# Essays in Environmental and AI Finance

Parvati Neelakantan

Student Number: 16200005


A dissertation submitted to University College Dublin in fulfilment of the requirements for the degree of Doctor of Philosophy in Banking and Finance

UCD Michael Smurfit Graduate Business School

Thesis Supervisor: Prof. Cal Muckley

Doctoral Studies Panel:

Prof. Thomas Conlon
Prof. Alessia Paccagnini
Prof. Valerio Poti

April 2022

## Dedication

*To my late grandmothers, Visalakshi and Rukmini, who continue to light my path.*

# Table of contents

# List of Tables

# List of Figures

# Acknowledgements

I am much grateful to my thesis supervisor, Professor Cal Muckley, for his unstinting support, able guidance, constant encouragement, and unfailing patience. I feel blessed to have worked with an incredible mentor who demystified the intricacies of research and helped me take my first steps into academia. His commitment to excellence, commendable work ethics, sense of propriety, timely and appropriate feedback motivated me to work diligently, think critically and become result-oriented. I also can't thank him enough for encouraging me to step out of my comfort zone and patiently helping me to overcome my inhibitions. It was truly a privilege to carry out my doctoral project under his nurturing supervision.

I am also indebted to Professor Antoine Dechezlepretre (London School of Economics; Organisation for Economic Co-operation and Development) for his guidance despite his intensely demanding schedule. Access to his expertise in Environmental Economics and Intellectual Property Rights has helped me immensely in writing this dissertation. Working along with him has been a great learning experience.

I would like to thank University College Dublin's faculty members, Professors Thomas Conlon, David Edelman, Benjamin Elsner, Alessia Paccagnini, and Ronan Powell for their invaluable contributions to this thesis.

During my PhD tenure, I was fortunate to seek guidance from Professors Ashwini Agrawal (London School of Economics), Wasim Ahmad (Department of Economic Sciences, Indian Institute of Technology Kanpur), Frank Barry (Trinity Business School, Trinity College Dublin), Donald Bergh (Daniels College of Business, University of Denver), Phelim Boyle (Lazaridis School of Business and Economics, Wilfrid Laurier University), Ron Giammarino (Sauder School of Business, The University of British Columbia), Paolo Guasoni (School of Mathematical Sciences, Dublin City University), John McConnell (Krannert Graduate School of Management, Purdue University), and Marti G. Subrahmanyam (Stern School of Business, New York University). I wish to thank them all for their insightful suggestions, encouragement, and inspiration.

I am also grateful to the funding organisations for their generous support which enabled me to conduct my research, attend training programmes and conferences. I am indebted to the Valuation and Risk Partnership (VAR), Science Foundation Ireland for offering me a fully funded PhD Scholarship, covering my conference expenses, and providing generous financial aid to attend intensive short-term training courses organised by the London School of Economics and Volatility Institute, NYU Shanghai. I would also like to gratefully acknowledge the financial support of Centre for Doctoral Research, University College Dublin for partly covering my training programme at London School of Economics and the Graduate Research Board, University College Dublin for providing me funding during the final stage of my doctoral studies.

I also truly appreciate the efforts of Dr. Na Li, Research Manager of VAR, for making my PhD journey smooth in every possible way. Additionally, I cannot even begin to explain how much Ms. Jane O'Mara, Manager of Centre for Doctoral Research, has meant to me during my PhD tenure. Her kindness, warmth, foresight, and amazing problem-solving ability made her the first person I turned to for advice on the administrative challenges that arose during the last four years. She has been a true friend and I will always cherish her quick wit, humour, and optimism.

I would like to thank my fellow PhD scholars at the School of Business and School of Economics for their insightful ideas and lively intellectual company. I would particularly like to mention Shivam Agarwal, Yuting Chen, Karen-Ann Dwyer, Eoin Flaherty, Yumeng Gao, Kushagra Jain, Yanan Lin, Sid McDonnell, Mark Regan, Xuanyu Yue, and Yang Zhao.

Further, I would like to thank my dear friends Caitriona, Fran, Harman, Kriti, Miriam, Sadhbh, and Sandra for their companionship, warmth, and affection. I truly appreciate their tremendous support in writing of my thesis.

Continents away from my family, I feel incredibly blessed to have met Ms. Rosemary St. Leger, Mr. Shamus Quinn, and Tadhg St. Leger Quinn. They were gracious enough to welcome a mere tenant as part of their family. I will always treasure the memory of the numerous delightful walks with Rosemary, hearing Shamus recount his enchanting sailing adventures, and Tadhg's spontaneous genial humour.

I would also like to express my gratitude to my brother and sister-in-law, Kailash Neelakantan and Radha Raghavan, for being so nurturing and dependable. They have been endlessly patient with my sulks and always nudged me towards clarity. Special thanks to my uncle, K.G. Raghavan, for his infectious enthusiasm and encouragement.

My biggest debt of gratitude goes to my parents, Gurumurthy Neelakantan and Rathna Neelakantan, for their unconditional love, support, and teaching the value of hard work by their example. They have always encouraged me to aim high and strive for things that I assumed I was incapable of. And this thesis would have never come about without their support and belief in me. Ma, Pa, thanks for everything.

To all the people who walked along with me on this journey,
Thanks a million!

# Abstract

Entitled "Essays in Environmental and AI Finance," this dissertation consists of three self-contained essays. The first essay avails of capital market price signals to assess the presence and magnitude of economic incentives for clean innovation relative to dirty innovation. Second essay examines the utility and ethics of incorporating national culture profiling in bank-level machine-learning informed alert models relating to financial malfeasance. And the third essay tests state-of-the-art model-agnostic explainable AI (XAI) methods to uncover algorithmic injustice in the bank lending space.

Essay 1 that seeks to bring new insights to the corporate environmental – financial performance debates examines how Tobin's Q is linked to 'clean' and 'dirty' innovation and innovation efficiency at the firm level. While clean innovation relates to patented technologies in areas such as renewable energy generation and electric cars, dirty innovation relates to fossil-based energy generation and combustion engines. A global patent data set covering over 15,000 firms across 12 countries helps uncover strong and robust evidence that the stock market recognizes the value of clean innovation and innovation efficiency and accords higher valuations to those firms that engage in successful clean research and development activities. The results are substantively invariant across innovation measurement, model specifications, estimators adopted, select sub-samples of firms and the United States and European patent offices.

Essay 2 examines the utility and ethics of incorporating national culture profiling in bank-level machine-learning informed alert models relating to financial malfeasance. On a globally significant financial institution, binary classifier type alert models are used to establish the utility of dimensions of national culture in formulating anti-money laundering predictions. For corporate (individual) accounts, Hofstede individuality (individuality, and national-level corruption perception and financial secrecy) scores of the country in which a customer is resident, or from which a wire is sent/received, are of paramount importance. When combined with extensive account and transaction data against an even proprietary institutional algorithm, national culture traits markedly enhance the models' predictive performances. Against a global standard, ethical implications of ascribing values to dimensions of national culture are examined. We posit an ethical framework for the use of national profiling in anti-fraud alert models.

Essay 3 provides evidence of the validity of Shapley model-agnostic explainable AI methods' on real-world datasets. This work contributes initial evidence on the usefulness of Global Shapley Value and Shapley-Lorenz methods, with respect to racial discrimination in lending. Using 157,269 loan applications from the Home Mortgage Disclosure Act data set in New York during 2017, it is confirmed that the methods reveal evidence of racial discrimination inherent in the predictions of a transparent logistic regression model. Thus explainable AI can enable financial institutions to select an opaque creditworthiness model which blends out-of-sample

performance with ethical considerations.

# Declaration

I hereby certify that this material submitted for assessment to the programme of study leading to the award of Doctor of Philosophy in Business Studies is entirely my own and has not been taken from the work of others, save to the extent that such work has been cited and acknowledged within the text of my work. I agree that the library may lend or copy the thesis upon request.

# Introduction and Overview

## 1.1 Introduction

One of the most pressing challenges of contemporary climate change policy concerns providing firms with the best incentives to redirect innovation away from fossil fuel (dirty) and towards low-carbon (clean) technologies. Prior studies highlight that certain policies can incentivize clean innovation, while discouraging dirty innovation (Calel and Dechezlepretre, 2016; Newell et al., 1999; Popp, 2002). In my dissertation, I investigate whether enough incentive exists to produce clean technologies. In other words, I examine whether a positive or negative incentive applies to clean innovation, given the plethora of factors governing a company's decision to produce clean or dirty technologies.

While there is strong evidence to suggest that R&D expenditure, in general, is linked positively to Tobin's Q (Griliches, 1981; Grandi et al., 2009), it does not necessarily follow that investment in the specific case of R&D to produce clean patents will likewise influence Tobin's Q. It is a moot empirical question whether a higher market valuation (higher Tobin's Q) follows in the case of clean vs dirty innovation. Further, prior literature foregrounds evidence of several mechanisms by which investment in environmental innovation (e.g., R&D expenditure on clean patents) can enhance or slow down a firm's financial performance. For instance, Fisher-Vanden and Thorburn (2011) and Jacobs et al. (2010) show that the market value of firms that voluntarily chose to produce clean technologies deteriorated. This, from the perspective of capital market participants, may be due to a compromised capital budget associated with clean innovation. On the other hand, clean innovation can also raise market value via mechanisms including attracting and retaining high quality employees (Dowell et al., 2000), avoiding regulatory penalties (Karpoff et al., 2005), and attracting ethical investors (Heinkel et al., 2001). As a result, it is ultimately an open empirical question whether clean innovation, by way of patents, impacts market evaluations and Tobin's q positively or not.

In this dissertation, I also examine whether national culture traits profiling can usefully inform a machine learning alert model to detect money laundering at a globally prominent financial institution. In light of recent literature on the role of culture in corporate misconduct and bank failure (Liu, 2016; Berger et al., 2021), I explore the relevance of several country-specific cultural and institution quality indices vis-à-vis modelling incidence of suspicious money movement within a financial institution.[1] As individuals may not always hold unbiased beliefs and

[1]The process of money laundering is a channel to legitimise dirty money (i.e., money generated from illegal activities) by integrating it into an established financial system for subsequent use without evoking suspicion. In facilitating the generation and disbursement of illicit proceeds from criminal activities, money laundering compounds the problem by paving the way for further financial illegal activity. Although difficult to measure, estimates for the total amount of money laundered worldwide range from 2-5% of global GDP (approximately $600 billion to $1.6 trillion).

can behave irrationally (Kim et al., 2016), the anticipated incentives and deterrents for misconduct and the anticipated likelihood of being held accountable for wrongdoing, can vary substantially across national cultures (Husted, 2000). The social normativity of national culture (Goodell, 2019), in particular, can influence misconduct among the customers of financial institutions. In my thesis, I provide practical implications for the financial services sector in terms of anti-money laundering compliance strategy.

Further, I test the validity of Global Shapley Value and Shapley-Lorenz model-agnostic explainable AI methods on a real-world finance dataset. Prior studies have tested Shapley value-based model-agnostic explainable methods' validity on simulated datasets (Štrumbelj and Kononenko, 2010; Štrumbelj and Kononenko, 2014; Aas et al., 2021). However, evidence for the methods' validity on real-world datasets, particularly in respect to impactful financial decisions is scant. Additionally, I note a paucity of studies applying model-agnostic explainable methods and, in particular, the Shapley Value methodology in the financial economics literature. Colombo and Pelagatti (2020) investigate the relative importance of variables in predicting movements in exchange rate models using partial dependence plots and permutation measure. Drehmann and Tarashev (2013) use the Shapley Value methodology to measure systemic importance of interconnected banks. Further, Tarashev et al. (2016) use the Shapley Value approach for risk attribution and to derive measures of banks' systemic importance. However, extant literature does not test usefulness of Shapley-based model-agnostic explainable methods in real-world datasets. In this thesis, I provide initial real-world evidence on the usefulness or otherwise of Global Shapley Value and Shapley-Lorenz methods in uncovering racial discrimination in the lending space. To the best of my knowledge, this study is perhaps the first of its kind to appropriately test the usefulness of the said methods in the bank lending space.

In the next Section, I specify the thesis's main research questions. In Section 1.3, I explain what motivates my research. In Section 1.4, I briefly discuss the data and methodologies employed in the thesis and its major findings. In Section 1.5, I provide an outline of the thesis's structure. In Section 1.6, I discuss the published work from the thesis. Further, in Section 1.7, I list the various institutions and conferences where my results have been presented, debated, and developed. Finally, in Section 1.8, I briefly summarise and conclude my discussion.

## 1.2 Main Research Questions

This dissertation through rigorous state-of-the-art statistical techniques addresses pertinent and timely research questions that carry enormous social impact. It addresses questions on climate change and ethical AI in the financial economics space. The three specific research questions that this thesis discusses are as follows.

The main research question addressed in Chapter 2 concerns whether a clean innovation premium exists consistent with the objective for a long-term de-carbonization of the international

economy. To resolve this question, I avail of a compelling litmus test consisting in the information content of equity market price signals. Hence, in Chapter 2, I investigate whether there is, from a market information assimilation perspective, an incentive for firms to pursue strategies of clean environmentally supportive innovation, as opposed to carbon-emitting dirty innovation activities.

While there is strong evidence that R&D expenditure is, in general, linked positively to Tobin's Q (Griliches, 1981; Grandi et al., 2009), it does not necessarily follow that investment in the specific case of R&D to produce clean patents will likewise influence Tobin's Q. Though the exclusive patent rights allow firms to control the production and distribution of their inventions and confers on them the right to a portion of the revenues of the competitors who use their technology, this does not guarantee the said firms higher market valuation (higher Tobin's Q) in the case of clean vs dirty innovations. Chapter 2 further highlights several mechanisms by which investment in environmental innovation (e.g., R&D expenditure on clean patents) can enhance or slow down a firm's financial performance. For instance, Fisher-Vanden and Thorburn (2011) and Jacobs et al. (2010) show that the market value of firms that voluntarily chose to produce clean technologies deteriorated. This may be due to a compromised capital budget, from the perspective of capital market participants, which is associated with clean innovation. On the other hand, clean innovation can also raise market value via mechanisms such as attracting and retaining high quality employees (Dowell et al., 2000), avoiding regulatory penalties (Karpoff et al., 2005), and attracting ethical investors (Heinkel et al., 2001). Thus, it is ultimately an open empirical question whether clean innovation, by way of patents, impacts market evaluations and Tobin's q positively at all.

In Chapter 3, I examine the utility and ethics of incorporating national culture profiling in bank-level machine-learning informed alert models relating to financial malfeasance. Specifically, I test to establish the utility of national culture traits informing a machine learning alert model for detecting money laundering at a globally prominent financial institution. National culture figures prominently in assessing qualities of ethics and discernment in business ethics research. This Chapter in strongly addressing business ethics research considers the use of national culture in machine-learning. It examines the timely and germane issue underlying the claim that latent racial and ethnic biases may inform instances of functional profiling or predictive models. I further assess the importance of national culture traits relative to customers' account and transaction traits. In so doing, this Chapter investigates if a banking customers' socio-cultural matrix inspires their predilections for committing money-laundering. This Chapter further provides the first description of the ethics associated with employing national culture profiles in machine-learning to counter money laundering.

Finally, Chapter 4 examines whether the feature importance in logistic regression predictive models as indicated by Global Shapley Value and Shapley-Lorenz model-agnostic explainable

AI methods align with evidence of feature importance in the underlying models, in the context of real-world financial services bank lending data. Scholarship confirms the validity of Shapley value-based model-agnostic explainable AI methods on simulated datasets (Štrumbelj and Kononenko, 2010; Štrumbelj and Kononenko, 2014; Aas et al., 2021).[2] However, evidence of their usefulness on real-world datasets is scant, particularly in respect to impactful financial decisions. The methodology adopted in Chapter 4 involves the estimation of tractable and transparent machine learning model, logistic regression, in mortgage lending data to discern the relative importance of predictive features. It then deploys Global Shapley Value and Shapley-Lorenz explainable AI methods to test if their insight concerning feature importance is in line with that of the logistic regression model. Finally, it examines whether these methods can enable financial institutions to select an opaque creditworthiness assessment model which blends out-of-sample performance with ethical considerations.

## 1.3 Motivation

The research questions raised in Chapter 2 are informed by several reports of Intergovernmental Panel on Climate Change (IPCC) which indicate that stabilizing global carbon emissions by 2050 will require a 60% reduction in the carbon intensity of global GDP compared with a business-as-usual scenario. Hence, the Chapter are motivated by whether a capital market incentive exists to decarbonise the international economy through a radical change in the mix of technologies that help produce and consume energy, rather than through energy-efficiency improvements of existing carbon-based technologies. From the perspective of an efficient markets argument, the capital market provides a summative signal of the complex combination of factors which provide incentives to innovate in radically new clean technologies or in improved dirty technologies.

Recent emphasis on clean technologies from fossil fuel-based innovations to curb carbon and other greenhouse gas emissions has inspired both theoretical and empirical research in this area. Using their microeconomic model, Acemoglu et al. (2012), have found that in the US energy sector the transition to clean technologies is likely to be delayed if fossil fuel-based technologies continue to prevail. In the same vein, Aghion et al. (2016) claim that firms in the auto industry are self-perpetuating and tend not to deviate from the type of innovation they are already invested in. However, they also note that production in innovation depends on aggregate location-based spill overs and that the firms facing higher tax-inclusive fuel prices may gravitate towards producing clean relative to dirty technologies. More pertinent, studies foreground evidence that firms may redirect innovation away from fossil fuel towards low carbon

[2]For instance, Štrumbelj and Kononenko (2010; 2014) use Shapley value-based feature importance measurements, via their approximation method, and show accurate results across various data generating processes. They use various learning algorithms such as decision trees, naïve bayes, support vector machines, multi-layer perceptron artificial neural networks, random forest, logistic regression and ADaBoost to evaluate and validate their approximation method. They further evaluate their method's usefulness on a real-world oncology dataset.

technologies, when faced with change in policies and energy prices. For instance, Calel and Dechezlepretre (2016) investigating the impact of the European Union Emissions Trading System, the largest carbon market in the world, on regulated companies discover that the policy caused regulated companies to increase patenting activity in low-carbon technology by 30%. Similarly, Newell et al. (1999) and Popp (2002) report a substantial increase in the production of energy-efficient technologies following increase in energy prices. However, a limitation of existing studies of directed technological change is that a multitude of drivers determine companies' decisions to conduct R&D activity in clean or dirty technologies. A complex medley of factors including the relative prices of production factors (Hicks, 1932b; Popp, 2002; Acemoglu et al., 2012), the quality of environmental policy instruments (Johnstone et al., 2010), the extent of market demand and a path-dependency in knowledge creation (Acemoglu et al., 2012; Aghion et al., 2016) can influence the prospective economic returns of clean and dirty innovation. Most important, many coexisting policies in a given jurisdiction - for example, carbon markets, fuel taxes, energy efficiency standards and renewable energy mandates - make it difficult to measure the overall stringency of environmental regulations faced by companies. An additional complexity consists in the expected realization of these policies and drivers which determine innovation decisions, rather than current observed realizations. However, these expectations are inevitably not directly observed and may vary markedly across firms. A major advantage of the approach adopted in Chapter 2 is that the stock market evaluation of patented innovation in clean and dirty technologies can reveal the market expectations with respect to the prospective economic performance of these investments which incorporate all their determinants, in particular from policies.

My motivation to study the research question raised in Chapter 3 stems from prior studies that associate national culture dimensions to financial misconduct (Liu, 2016; DeBacker et al., 2015; Bame-Aldred et al., 2013); quality of ethical behaviour and perception (Armstrong, 1996; Davis and Ruhe, 2003; Getz and Volkema, 2001; Vitell et al., 1993; Volkema, 2004); and finance-related behaviour (Chui et al., 2010; Lievenbrück and Schmid, 2014; Aggarwal and Goodell, 2009; Chui et al., 2002; Shao et al., 2013; Aggarwal and Goodell, 2013). Given the breadth of scholarship associating national culture with behaviour in business, there are ample reasons to investigate whether national culture impacts an individual's or corporation's predilection for bank fraud. In Chapter 3, I test to establish the utility of national culture traits informing a machine learning alert model for detecting money laundering at a globally prominent financial institution. More pointedly, this study examines if a banking customers' socio-cultural matrix informs their predilections for committing financial misconduct, namely, money-laundering. This Chapter addresses the issue behind the claim that latent racial and ethnic biases may inform instances of functional profiling or predictive models. I also construct and critique an ethical framework in respect to the employment of national profiling in anti-fraud alert models.

Finally, the motivation for Chapter 4 is drawn from the studies at the Bank of England, Finan-

cial Conduct Authority, as well as at the European Banking Authority that highlight, notwithstanding the impressive predictive performances evident in the deployment of machine learning models in banking, the opacity of complex machine learning models comprises a significant impediment to their implementation. Further, regulators and various national agencies in the US (USACM, 2017; OSTP Report, 2016) and Europe (European Commission, 2019; France, 2018a and 2018b) are increasingly recognising the importance of algorithmic transparency and accountability. They encourage the use of Machine Learning models that ensure high predictive performance, even while informing the interpretability of models. For instance, the European Commission emphasises the importance of research in explainable AI systems to render transparent and accountable high performance machine learning models with a view to ensuring the protection of customer rights. Although black-box models may yield impressive predictive performances, their obfuscating internal logic may inadvertently perpetuate biases leading to prevention of detection and mitigation of discrimination. In revealing the importance of features that determine the machine learning models' decisions, the state-of-the-art explainable model-agnostic methods can uncover algorithmic biases and, thereby allow institutions to employ "fairness" techniques for rectifying the error. Hence, the explainable AI tools revealing whether the algorithms are fair ensures the management and regulators' trust in the models, leverage the use of complex models to yield profitable and fair outcomes besides helping firms conceal their intellectual property. Chapter 4 delivers practitioner-oriented tests and demonstrations on the usefulness of Shapley measures that render opaque but accurate machine learning models useful, in line with the spirit of regulatory supervision governing algorithmic bias and model accountability. Thus, this Chapter provides real-world evidence on the usefulness or otherwise of the explainable AI techniques in uncovering racial discrimination in the bank lending space.

## 1.4 Data, Methodology, and Major Findings

The dataset employed in Chapter 2 includes firm-level data from Worldscope and Datastream; and patent-level data from the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office (EPO). The firm-level data sourced from the Worldscope database encompasses financial and accounting information of listed firms drawn across forty countries during the years 1995-2012. Specifically, the original sampled dataset comprises 47,420 firms in 40 countries. However, after cleaning this dataset the final firm-count stands at 25,255. Next, the stock-price data for these listed firms is collected from Datastream. To match the firm-level data with patent data, I employ Bureau van Dijk's matching algorithm provided under the "IP" bundle of the Orbis database.[3] While matching the name, this algorithm also matches geographical information, which is sourced from patent data (country, address, etc). Furthermore, the matching process undergoes extensive manual cleaning to ensure a firm matches with its

[3]Bureau van Dijk owns Orbis.

patents accurately.[4]

I primarily focus on the patents and citations published by the United States Patent and Trademark Office (USPTO), in line with existing studies; however, for robustness I also analyse the patents and citations published by the EPO. The baseline empirical analysis focuses on a sample of USPTO published patents and citations filed by 15,217 firms belonging to the top 12 leader countries in clean innovation during 1995-2012.[5]

After matching the firm-level data with patent-level data, I create the innovation variables. While these variables are inspired by prior literature (Deng et al., 1999; Chan et al., 2001; Gu, 2005; Hirshleifer et al., 2013), the chief novelty of my study consists in disaggregating these into 'clean,' 'dirty,' and 'other' components. To determine the expected economic performance of 'clean' and 'dirty' investment, I adapt a firm-level market-value function (Griliches, 1981; Hall et al., 2005; Hall and Oriani, 2006) and Fama-MacBeth regressions (Fama and MacBeth, 1973). In the light of these models, I draw inferences on how the innovation variables influence the firm's Tobin's Q.

Chapter 2 reports evidence that 'clean' innovation typically yields positive associations with Tobin's Q. While this result is economically significant, the capital market ascribes no (or a negative) market value influence on 'dirty' innovation. The relative Tobin's Q association of 'clean' vis-a-vis 'dirty' innovation is significant and economically important across innovation measurements. Further, I adopt a wide variety of complementary and state-of-the-art testing procedures to investigate whether clean innovation is associated with firm value. Across the conducted tests, I show evidence of the importance of clean innovation (but not dirty innovation) for the equity market's indication of firm value.

Chapter 3 employs a major global financial institution's large proprietary dataset containing cross-border wire transactions made during 2009-2018. Those wires that the institution's designated investigative team flagged as 'suspicious activities' can be regarded as precursors to money laundering. I further collate the novel proprietorial customer and account level cross-border wire transfer bank client data with country-specific culture (Hofstede's cultural dimensions) and institution quality indices (Corruption Perception Index; Financial Secrecy Index). Further, the proprietorial dataset provides a clearly labelled response variable (Issue Case). I, therefore, employ supervised learning techniques such as logistic regressions, random forest, gradient boosted machines, and support vector machines to detect money-laundering at the financial institution. Employing the said machine learning techniques together with corrections for data imbalance, the results reflect the strength of national culture dimensions in formulating

[4]Note that in using the harmonized version of patent applicant names from PATSTAT for carrying out the firm-patent matching, Orbis largely mitigates any improper matching of financial data with patent data.

[5]The top 12 clean innovation producing countries in descending order are: Japan, USA, Korea, Germany, Taiwan, France, Denmark, Netherlands, Canada, Sweden, Finland, and Great Britain. Patenting at the USPTO in clean and dirty technologies becomes miniscule beyond these top 12 countries.

anti-money laundering (AML) predictions. Further, the introduction of these variables complements the institution's own account- and transaction-level data, considering that the inclusion of these predictors as an added layer of attributes enhances the performance of the models. These findings provide practical implications for the financial services sector in terms of AML compliance and prevention strategy. Confirming the conduciveness of machine learning in incorporating national culture, the findings also contribute to the extensive literature that ascribes values to ethicality and discernment constituting distinct national traits.

In Chapter 4, I test Global Shapley Value and Shapley-Lorenz model-agnostic explainable AI (XAI) techniques for interpreting and understanding a machine learning model's internal logic that determines an applicant's creditworthiness. Derived from Game Theory, Shapley value-based model-agnostic XAI methods explain machine learning models' predictions, by assuming, for each data point, that each feature value is a "player" in a game with prediction being the payout. Theoretically, Shapley values are the "fair" distribution of the payout among features. In comparison to other approaches, such as partial dependence plots and permutation methods, Shapley values fairly distribute the difference between the prediction and the average prediction among the features (Molnar, 2020). As a result, they rank as insightful methods to shed light on the predictive machine learning models' internal logic.

Prior literature finds evidence of racial discrimination in both in-person and algorithmic lending in the US (Black et al., 1978; Munnell et al., 1996; Blanchflower et al., 2003; Butler et al., 2020; Bartlett et al., 2022). Thus, I test if Global Shapley Value and Shapley-Lorenz XAI methods can uncover algorithmic injustice in the bank lending space. Further, 157,269 loan applications from Home Mortgage Disclosure Act's (HMDA) website made in New York during 2017 is examined. I first deploy a logistic regression model and show evidence consistent with racial discrimination. I then test if the said XAI methods give insight consistent with the logistic regression model. Accordingly, I find that these XAI methods establish the prevalence of racial discrimination as a paramount factor. In revealing that the XAI methods uncover racial discrimination, the analysis confirms their validity in respect to the logistic regression model, and in real-world datasets. This Chapter also shows how financial institutions can derive accurate and accountable decisions, in the context of racial discrimination and opaque credit-worthiness models.

## 1.5 Thesis Structure

The thesis consists of three essays that address important social questions on climate change and ethical AI in the financial economics space. Chapter 2 that deals with environmental finance address whether an economic incentive obtains for firms to pursue strategies of clean environmentally supportive innovation, as opposed to carbon-emitting dirty innovation. Chapters 3 and 4 that discuss financial data science assess the utility and ethics of incorporating

national culture profiling in bank-level machine-learning informed alert models relating to financial malfeasance and tests state-of-the-art explainable AI techniques to uncover algorithmic injustice in the bank lending space, respectively. The final chapter neatly sums up the bottom-line conclusions.

Chapter 2 highlights several mechanisms by which investment in environmental innovation can enhance or slow down a firm's financial performance. While I outline the mechanisms that can account for a clean innovation premium/discount, my purpose is not to empirically test an individual mechanism which can account for the principal finding. Rather, I elicit from the equity market data an evaluation of clean vis-à-vis dirty innovation. While the question I address is important, its resolution is far from straightforward. Leaving the identification of a mechanism that accounts for the principal finding of a clean innovation premium to future research, I seek to resolve the raised question by availing of a compelling litmus test that consists in the information content of equity market price signals. To meaningfully address the research question, I construct innovation variables. The chief novelty of my work consists in disaggregating these variables into 'clean,' 'dirty,' and 'other' components. Thus, I provide detailed description of variable construction in the chapter. Further, I discuss the methodology employed to perform a range of well-motivated empirical tests. This is followed by a discussion of my findings. Finally, I summarise the findings and discuss the direction of related future research in the concluding section.

Chapter 3 establishes through binary classification alert models the utility of national culture in formulating anti-money laundering predictions in a globally prominent financial institution. Accordingly, I commence by providing the rationale for examining whether national culture can inform anti-money laundering alert models. I then discuss the proprietary dataset, country-specific culture, and institution quality indices from which I have drawn my predictors/features. After a detailed description of the various data resampling methods used in the chapter for meaningfully sourcing information from the data, I discuss the machine learning methodologies, performance evaluation metrics, and feature importance metrics the study employs. Finally, I present the empirical findings and provide a framework for evaluating the ethics of machine learning prediction and alert models.

Chapter 4 delivers practitioner-oriented tests and demonstrations on the usefulness of Global Shapley Value and Shapley-Lorenz measures in rendering opaque but accurate machine learning models, in line with the spirit of regulatory supervision informing algorithmic bias and model accountability. While the advent of AI has meant faster and historically accurate lending decisions, its models often fail to enhance the decisions' accountability. So, regulators and various national agencies in the US (USACM, 2017; OSTP Report, 2016) and Europe (European Commission, 2019; France, 2018a and 2018b) have begun to stress the value of algorithmic transparency and accountability. Therefore, I begin the chapter by discussing the evidence that

prior studies have discovered of racial discrimination in both in-person and algorithmic lending in the US. I then discuss the data and variables employed in the chapter to model an applicant's creditworthiness. This is followed by a detailed discussion of the Global Shapley Value and Shapley Lorenz explainable methods and presentation of the empirical findings.

In chapter 5, I discuss the main findings and limitations of the thesis. I also identify avenues for future research.

## 1.6 Published Work from this Thesis

The contributions reported in Chapter 2 are presented in Dechezleprêtre, Muckley, and Neelakantan (2021), published in *The European Journal of Finance* and Dechezleprêtre, Muckley, and Neelakantan (2021), published in A.B. Dorsman, K.B. Atici, A. Ulucan, M.B. Karan (eds), "Applied Operations Research and Financial Modelling in Energy Practical Applications and Implications."

## 1.7 Conference Presentations

The findings of Chapter 2 were presented at the Financial Data Science and Econometrics Workshop (Loughborough, United Kingdom, September 2018), International Conference on FinTech & Data Science (Dublin, Ireland, September 2019), 18th International Conference on Credit Risk Evaluation Designed for Institutional Targeting (CREDIT) in Finance (Venice, Italy, September 2019), and during the International Economics and Finance session of TBS AIB Paper Development Workshop (Dublin, Ireland, October 2019) organized by Trinity Business School, Trinity College Dublin. They were also presented at the UCD Graduate Research Student Symposium (May 2019), VAR (Valuation and Risk) Research Day (May 2019), UCD College of Business PhD Symposium (September 2019), VAR Research Day (December 2019) held at University College Dublin, VAR Research Day (May 2020), VAR Research Day (December 2020), and 8th Multinational Energy and Value Conference (Leuven, Belgium, May 2021).

The contributions reported in Chapter 3 were presented at VAR Research Day (June 2021) held at University College Dublin and 'Women in FinTech' Conference (September 2021) organized by FinTech and AI Cost Action (CA 19130), Brussels, Belgium. The Chapter has also been accepted for presentation at the 31st European Financial Management Association (conference to be held at Bio-Medico University, Rome, Italy, during June 29– July 2, 2022), 35th Annual Irish Economic Association (organized by University of Limerick and to be held during 5-6 May 2022), World Finance Conference (organized by School of Management and Economics, University of Turin and to be held during 1-3 August 2022), 4th International Conference on Financial Markets and Corporate Finance (organized by Indian Institute of Technology Bombay and to be held during 7-9 July 2022), The Finance Symposium 2022, Chania, Crete, Greece

(29-31 July 2022), and 2022 IFABS Conference (hosted by the University of Naples, Naples, Italy, during 7-9 September, 2022).

Finally, the findings reported in Chapter 4 were presented at 'Women in FinTech' Conference (September 2021) organized by FinTech and AI Cost Action (CA 19130), Brussels, Belgium and 3rd International Conference on Digital, Innovation, Entrepreneurship & Financing (December 2021) organized by INSEEC School of Business and Economics, John Molson School of Business, Concordia University, and School of Economics, Jilin University. The Chapter has also been accepted for presentation at the 3rd Irish Academy of Finance (IAF) Conference (Dublin, Ireland, May 2022), Economics of Financial Technology Conference (University of Edinburgh, May 2022), the 11th International Conference of the Financial Engineering and Banking Society (jointly organized by the Portsmouth Business School, University of Portsmouth, UK and the Montpellier Business School, France, during June 2022), 29th Annual Global Finance Conference (organized by University of Minho, Braga, Portugal, during June 2022), 4th International Conference on Financial Markets and Corporate Finance (organized by Indian Institute of Technology Bombay and to be held during 7-9 July 2022), and 2022 IFABS Conference (hosted by the University of Naples, Naples, Italy, during 7-9 September, 2022)

## 1.8 Summary and Conclusions

I begin this chapter by foregrounding the thesis's main research questions and my motivation for pursuing them. Further, after briefly outlining my data and methodology, I discuss the thesis's major findings. I then present the thesis structure, chapter by chapter. This is followed by a discussion of my published work from the thesis and the various venues where my research work was presented and critiqued.

# Is firm-level clean or dirty innovation valued more?

**Abstract**

We examine how Tobin's Q is linked to 'clean' and 'dirty' innovation and innovation efficiency at the firm level. Clean innovation relates to patented technologies in areas such as renewable energy generation and electric cars, whereas dirty innovation relates to fossil-based energy generation and combustion engines. We use a global patent data set, covering over 15,000 firms across 12 countries. We find strong and robust evidence that the stock market recognizes the value of clean innovation and innovation efficiency and accords higher valuations to those firms that engage in successful clean research and development activities. The results are substantively invariant across innovation measurement, model specifications, estimators adopted, select sub-samples of firms and United States and European patent offices.

**JEL Classification: G35, G32, C58**

**Keywords:** Innovation, research and development, patents, citations, clean technology, dirty technology, market value

## 2.1 Introduction

According to an Assessment Report by the Intergovernmental Panel on Climate Change, stabilising global carbon emissions in 2050 requires a 60% reduction in the carbon intensity of global GDP compared with a business-as-usual scenario (IPCC, 2014). In order to achieve a decarbonisation of the economy, while meeting growing global energy demands, the world needs to implement a radical change in the mix of technologies used to produce and consume energy. This, in turn, requires massive investments in research and development activities. For this reason, one of the most pressing challenges for climate change policies today is to ensure, in the context of multiple market failures associated with environmental externalities and R&D provision (Jaffe et al., 2005), that there is an adequate economic incentive for firms to redirect innovation away from fossil fuel ('dirty') and towards low-carbon ('clean') technologies. In this paper, we avail of capital market price signals to assess the presence and magnitude of economic incentives for clean innovation relative to dirty innovation. We examine whether firms conducting clean innovation trade at a premium or a discount relative to firms which conduct dirty innovation.

Understanding the determinants of clean technological change is a lively research area, both on the theoretical (Acemoglu et al., 2012) and on the empirical side (Aghion et al., 2016). Several studies have shown evidence that firms redirect innovation away from fossil fuel towards low-carbon technologies when faced with a change in policies or market conditions. For instance, Calel and Dechezlepretre (2016) investigate the impact of the European Union Emissions Trading System - the largest carbon market in the world - on regulated companies using a matching

method and report that the policy caused regulated companies to increase patenting activity in low-carbon technology by 30%. Similarly, Newell et al. (1999) and Popp (2002) report a substantial increase in the production of energy-efficient technologies following an increase in energy prices.

However, a limitation of existing studies of induced technological change towards clean innovation is that a multitude of drivers can determine companies' decisions to conduct R&D activity. These drivers include the relative prices of production factors (Hicks, 1932b; Popp, 2002; Acemoglu et al., 2012) but also the quality of environmental policy instruments (Johnstone et al., 2010) and the extent of a path-dependency in knowledge creation and market demand (Acemoglu et al., 2012; Aghion et al., 2016), which can all influence the prospective economic returns of clean and dirty innovation. Critically, a variety of policies and drivers can coexist in a given jurisdiction - for example, carbon markets, fuel taxes, energy efficiency standards and renewable energy mandates - making it difficult to measure the overall impact of these policies and drivers taken together or considered in isolation. An additional complication is that it is the expected realization of these policies and drivers which determine innovation decisions, rather than current observed realizations. But these expectations are inevitably not directly observed and may vary markedly across firms. A major advantage of our approach, relative to extant studies, is that the stock market evaluation of patented innovation in clean and dirty technologies can reveal market expectations with respect to the prospective economic performance of these complex investments.

Our analysis avails of a global firm-level patent data set, covering 15,217 firms across 12 countries. Our patent data are drawn from the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office (EPO). Our database reports the name of patent applicants, which allows us to match clean and dirty patents with distinct patent holders. The global nature of the database means that we can test our hypothesis on several measures of patenting activity, including patents taken out in the world's major patents offices such as the United States Patents and Trademark Office (USPTO) or the European Patent Office (EPO), irrespective of the jurisdiction of the innovating firm. Our data also includes information on patent citations, allowing us to address the well-known issue of heterogeneity in patent value. We associate 'dirty' innovation with fossil-based energy generation and ground transportation, and 'clean' innovation with renewable energy generation, electric vehicles and energy efficiency technologies in the buildings sector. The clean and dirty innovation categories allow us to, specifically, develop and study insightful dis-aggregated versions of well-known innovation productivity (Chan et al., 2001; Deng et al., 1999; Gu, 2005) and efficiency variables (Hirshleifer et al., 2013). We primarily study the patents and citations that are published by the USPTO, however for robustness we also conduct our analysis to the patents and citations published by the European Patent Office (EPO).

We first verify, in our sample, the capital market value accorded to generic innovation productivity (Deng et al., 1999; Chan et al., 2001) and innovation efficiency (Hirshleifer et al., 2013). This work serves to extend, in the international arena, the non-linear least squares regression model findings in Hall et al. (2005).[6] To determine if there is an economic incentive for firms to direct innovation away from fossil fuel ('dirty') and towards low-carbon ('clean') technologies, we regress firm-level Tobin's Q on firm-level clean and dirty innovation, together with innovation in other technologies. To ascertain the expected economic performance of 'clean' and 'dirty' investment activities, we, specifically, follow Hall et al. (2005) and adopt a firm's intangible stock of knowledge function. We dis-aggregate innovation productivity measures and innovation efficiency measures that are similar to those used in Deng et al. (1999) and Hirshleifer et al. (2013) to account for 'clean' and 'dirty' innovation production and efficiency, respectively.

Our main findings are as follows. Consistent with the view that the capital market evaluates clean innovation positively, we find that an additional clean patent, per million dollars of book value, is associated with an increment of 3.77% in Tobin's Q. We also find that generating a citation on a clean patent, per million dollars of book value, is associated with an increment of 1.27% in Tobin's Q. We also note that the comparable efficiency of R&D investments, in generating dirty patents, reduces the market value of the firm to the tune of 0.97% of its economic value. Our main finding is, thus, that 'clean' innovation is associated with an economically important and positive Tobin's Q relation, especially relative to the inferred association with dirty innovation.

We implement a series of robustness tests. These checks are based on a variety of dimensions: (i) we test, following Hirshleifer et al. (2013), if the findings are invariant to an alternative estimator, the Fama-Macbeth two-step regression estimator (Fama and MacBeth, 1973), (ii) we test if the results are robust to examining only those firms which conduct both clean and dirty innovation, (iii) we test if the results can be accounted for by including emerging technology innovation in our main regression equations, (iv) we check the sensitivity of the results to including a range of firm traits from the accounting based asset pricing literature (Ohlson, 1989,9; Hirshleifer et al., 2013), (v) we conduct a Heckman two-stage analysis (Heckman, 1979) to account for sample selection concerns, (vi) we test if our main findings hold when we examine European patents, as opposed to United States patents. Our main findings are substantively unchanged across all these tests.

Our paper relates to the extensive literature that links firm-level environmental performance with its financial performance. Earlier papers including Gupta and Goldar (2005) show that capital markets can create financial and reputational incentives for pollution control in both de-

[6]The initial findings corroborate a large body of research which provides compelling evidence that the patent productivity of R&D and the citations received by these patents have a statistically and economically significant positive impact on firms' market value (e.g. Griliches (1981), Chan et al. (2001) and Eberhart et al. (2004)).

veloped and emerging market economies (see also Hamilton (1995) and Dasgupta et al. (2001)). More recent papers such as that of Guenster et al. (2011) show that eco-efficiency relates positively to operating performance and market value (see also, Ziegler et al. (2007) and Von Arx and Ziegler (2014)). Prior studies, however, suffer from several problems including small samples and the lack of objective environmental performance criteria. We do not rely on subjective analysis to characterize environmental performance. Instead, we study the documented environmental patenting activity and the efficiency of this patenting activity of publicly traded firms around the world. In addition, this prior literature, unlike our paper, does not look at the critically important performance criterion of environmentally friendly patented innovation (IPCC 2014), with a view to improving the mix of technologies used to produce and consume energy. It does not, hence, examine whether this type of environmental performance can be related to financial performance and capital market values.

The remainder of the paper is organized as follows. Section 2 presents a discussion of possible mechanisms which can inter-relate market valuations and environmentally coherent innovation. Section 3 presents our data sources and characterizes our sample. Section 4 presents our econometric methodology. Section 5 presents our results and robustness tests. Section 6 concludes.

## 2.2 Theoretical background: Market evaluation of innovation and 'green' business decisions

Our point of departure is the well-established notion that stock markets can provide useful information on the value and expected performance of R&D investments (Griliches, 1981; Chan et al., 2001; Eberhart et al., 2004; Hall et al., 2005; Hirshleifer et al., 2013).[7] Assuming efficient capital markets, traded security prices can provide an unbiased estimate of the present value of discounted future cash flows. There exists, however, significant differences in the market value of R&D investments across time, sectors and countries (Grandi et al., 2009). What we examine in this paper, which has not been studied previously, is whether clean firm-level innovation productivity and efficiency are valued in capital markets around the world, in particular compared to dirty innovation productivity and efficiency. The literature identifies two potentially countervailing outcomes, which can prevail, between investments in environmental innovation and financial performance.

[7] As the returns to R&D investments will typically accrue over a number of years, stock prices or market value should provide, given market information efficiency arguments, useful information on their expected future benefits. Empirical studies analysing the relationship between R&D investments and market value typically model the market value relative to tangible assets (Tobin's Q) as a function of intangible assets (R&D capital), among other firm value determining variables, and show that the R&D-market value relationship is consistently positive (Ballardini et al., 2005).

### 2.2.1 Clean innovation and positive stock market evaluation

Low-carbon and more generally environmental innovation by firms can be evaluated positively in the capital market as it can increase expected firm-level cash-flows (revenues less costs) and/or reduce the risk of these cash flows. There is a variety of potential mechanisms which can link firm-level environmental innovation and financial performance. Due to the plethora of emissions trading systems, climate and energy policies around the world (Ellerman et al., 2014), such innovation not only has generic research and development expenditure implications for future firm operating cash flows and risks (Hall, 2000; Czarnitzki et al., 2006). It also reflects recipient firms' expected environmental taxes and subsidies and financial penalties for environmental policy violations.

First, to the extent that environmental innovation is a measure of environmental performance, investors can link pro-active environmental innovation to lower firm risk. For instance, environmental performance can proxy for (i) high-skilled management (Bowman and Haire, 1975) and labour conditions at the firm and thus the firm's capacity to attract high-quality employees (Turban and Greening, 1997) and increasing employee morale and productivity (Dowell et al., 2000); (ii) operational efficiency (Porter and Van der Linde, 1995); and (iii) sales benefits in existing markets (Klassen and McLaughlin, 1996) and in new markets (Porter and Van der Linde, 1995) due to improved corporate and brand reputation with regulators, employees and the public (Corbett and Muthulingam, 2008; Russo and Fouts, 1997). More generally, (iv) environmental innovation can be regarded as a less risky investment (Narver, 1971; Shane and Spicer, 1983; Spicer, 1978). There is also evidence that firms with high commitments towards corporate social responsibility offer lower wage and enjoy higher employee productivity due to better recruitment, higher intrinsic motivation (many employees prefer a socially responsible employer and will accept a lower wage to achieve this), and a more effort-promoting corporate culture (Nyborg and Zhang, 2013; Brekke and Nyborg, 2008).

It is also possible that the life-cycle of the technology sector of a clean patent can account for it being associated with a positive stock market evaluation.[8] Essentially, early stage life-cycle technology can be associated with potential for high growth albeit also high risk. If initially assets are valued above their replacement cost, competition in the marketplace will erode this mark-up over time (Tobin, 1969). Depending on the shape of this trajectory, innovation at a mature stage (e.g., internal combustion engines) will typically be valued less, relative to replacement cost, than innovation in relation to new technologies (e.g., energy generation through renewable energy sources). In a similar vein, this life-cycle argument can lead to smaller effects of incremental patenting on Tobin's q for a given technology over time (i.e., radical innovations are likely to precede incremental innovations in time). As a result, effects on Tobin's q for new technologies can be expected to be greater than for existing technologies that are in the refine-

[8] We thank an anonymous reviewer for raising this point.

ment phase of their life cycle, and are facing stronger competition. A life-cycle mechanism can potentially account for a clean innovation premium.

Third, climate change innovation can serve to mitigate risks of losses from crises or new regulation[9] (Reinhardt, 1999) and prevent expenses due to lawsuits and legal settlements (Karpoff et al., 2005). Investors can, hence, assign a lower discount rate to firms which are high environmental performers which would accord the firm a higher market value (and lower expected stock returns). Finally, climate change innovation can attract funds from ethical investors who can prefer firms with good track records of environmental performance (Heinkel et al., 2001). This interest on the part of ethical investment funds can reduce the cost of capital for the firm when it seeks to raise finance in the capital markets.

### 2.2.2 Clean innovation and negative stock market evaluation

To the contrary, it is also possible that corporate investment in environmental innovation can deteriorate a firm's financial performance (Walley and Whitehead, 1994; Palmer et al., 1995). Climate change innovation can also, thus, be associated with a negative stock market valuation impact. Fisher-Vanden and Thorburn (2011) and Jacobs et al. (2010) show that emissions reductions can be associated with significant negative market reactions. In particular, the stock market may respond negatively to such innovation due to the possibility that the capital budget of the firm is deteriorated by such investment. For instance, it may be interpreted by participants in the capital market that pertinent environmental legislation is binding at present or in the future. Environmental subsidies which are sought or the avoidance of financial penalties in respect to the emission of pollutants, which has motivated the environmental patenting activity, can also be ascribed a lower probability by capital market participants, than by firm management.

Two additional results, from the broad empirical R&D and market valuation literature, which can bias our inferences away from a clean innovation premium, should be highlighted. First, firms' market share positively impacts on the valuation of R&D (Blundell et al., 1999), and firms conducting 'dirty' innovation are typically large incumbents, while firms engaged in clean innovation are more likely to be new entrants. New firms are often the vehicle through which radical, game-changing innovations enter the market. Our sample of listed firms is over-representative of large firms, but even within listed firms, clean innovators might be smaller than dirty innovators. Second, a decreasing relationship between market uncertainty and the valuation of R&D investments has been observed (Oriani and Sobrero, 2008). Since the demand for clean innovation fundamentally depends upon environmental policies, which are inherently uncertain, this could lower the premium associated with pursuing environmental clean R&D investments.

[9]Calel and Dechezlepretre (2016) show that the European Union Emission's Trading System has had a quick causal impact on technological change in the form of new patenting activity.

## 2.3 Data and Variables

This section presents our sample of firm and patenting data, including a discussion of clean and dirty patent categories. It also presents our key variables of interest: Tobin's Q, innovation productivity and efficiency variables and control variables. Finally, it presents descriptive statistics in respect to the evolution of clean and dirty innovation globally.

### 2.3.1 Our sample of firms

Our sample of firms is obtained from the Worldscope Database, which presents information on the largest firms internationally. The original sampled data comprises 47,420 listed firms in 40 countries. From the original sample of firms, we eliminate firms for which the ISIN No. is missing, and we retain firms in the home market where the ISIN No. is the same for two firms in two different markets. Next, we drop firms with negative total assets, market capitalization or common cash dividend paid. We also drop firms for which we have less than 5 consecutive firm-year observations between 1995 and 2012 across a subset of firm-level variables - year-end market capitalisation, capital expenditure, and earnings before interest, tax and amortisation. The final firm-count is 25,255 firms from Worldscope.

### 2.3.2 The PATSTAT database

We use patent data to identify innovation in clean and dirty technologies. To construct our innovation variables, we have drawn data from the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office. PATSTAT is the largest international patent database, including all of the major offices such as the United States Patent and Trademark office (USPTO) and the European patent office. In PATSTAT, patent documents are categorized according to the new Cooperative Patent Classification system (CPC), the International Patent Classification (IPC) and national classification systems. For each patent we know at which date it was filed (the application date), when it was first published (the publication date) and, if it was ever granted by the patent office, when the granted patent was published. In our study, we focus on patent publication date as it is reasonable to expect that capital market participants will become aware of the new patents at this date.

The use of patent data has gained popularity in the recent empirical literature. An advantage of patent data is that they focus on outputs of the inventive process (Griliches, 1990). Furthermore, they provide a wealth of information on the nature of the invention and the applicant. Most importantly, they can be disaggregated to specific technological areas.

Patents also suffer from a number of limitations. The first limitation is that for protecting innovations, patents are only one of several means, along with lead-time, industrial secrecy, or purposefully complex specifications (Cohen et al. 2000; Frietsch and Schmoch 2006). However, a large fraction of the most economically significant innovations appear to have been

patented (Dernis et al., 2001). Moreover, in several sectors of which many clean and dirty technologies originate, such as automotive or special purpose machinery, patents are perceived as an effective means of protection against imitation (Cohen et al., 2000).[10] A second limitation is that the propensity to patent (e.g. the number of patents filed per USD of R&D) differs across industries and jurisdictions, making it difficult to use patent metrics for comparisons across sectors and countries. This problem can be alleviated in the econometric analysis by including industry and country fixed effects. Time fixed effects control for changes in the propensity to patent across time. A final problem is that patent values are highly heterogeneous, with most patents having a low valuation (Griliches, 1998). This problem is partly addressed by invoking the law of large numbers, since our large dataset (over 15,000 companies across 12 countries) enables us to analyse average differences in the association between patenting and Tobin's Q across technologies. In addition, we employ citation-adjusted patent counts in our models. It is widely accepted that citations received by patents are an indication of the economic significance of an innovation (Harhoff et al., 2003).

Our database in providing the identity of the patent applicants also facilitates matching clean and dirty patents with distinct patent applicants.[11] Our analysis focuses on a sample of published patents and citations, for listed firms for which we observe firm traits, filed by 15,217 firms belonging to the top 12 country leaders in clean innovation[12] over the period 1995-2012. We primarily study the patents and citations that are published by the USPTO, however for robustness we also conduct our analysis to the patents and citations published by the European Patent Office (EPO).

### 2.3.3 Clean and dirty patent categories

Our selection of patent classification codes for clean technologies relies on previous work by the OECD Environment Directorate.[13] We examine areas of clean patenting activity related to energy generation from renewable and non-fossil sources (wind, solar, hydro, marine, biomass, geothermal and energy from waste), combustion technologies with mitigation potential (for example combined heat and power), other technologies with potential contribution to emissions mitigation (in particular energy storage), electric and hybrid vehicles and energy conservation in buildings. We refer to these areas as climate change mitigation innovation or in short 'clean'

[10] Cohen et al. (2000) conducted a survey questionnaire administered to 1,478 RD labs in the U.S. manufacturing sector. They rank sectors according to how effective patents are considered as a means of protection against imitation, and find that the top three industries according to this criterion are medical equipment and drugs, special purpose machinery and automobile.

[11] To link patent applicants with firms in Worldscope, we use the link provided by Bureau van Dijk's Orbis database in its "IP" bundle, to which we have access through a commercial license. The matching algorithm is based not only on name matching but also on geographical information available from patent data (country, address, etc) as well as on extensive manual cleaning.

[12] The top 12 clean innovation producing countries in descending order are: Japan, USA, Korea, Germany, Taiwan, France, Denmark, Netherlands, Canada, Sweden, Finland and Great Britain.

[13] See www.oecd/environment/innovation

innovation. The patent classification codes used to extract clean patents from the database is presented in Table A1 in the Internet Appendix A.

Our selection of patent classification codes for dirty technologies relies on Noailly and Smeets (2015) for electricity generation technologies and on Aghion et al. (2016) for the automobile industry. Our dirty environmental innovation pertains to IPC codes in different technological classes, including steam engine plants, gas turbine plants, combustion engines, steam generation, combustion apparatus and furnaces. The patent classification codes used to extract dirty patents from the database are presented in Table A2 in the Internet Appendix A.

### 2.3.4 Key variables of interest and control variables

Our dependent variable, Tobin's Q, and independent variables, innovation productivity and efficiency variables, as well as control variables (i.e. firm trait variables) are described in this sub-section. Concise definitions are provided in Table 1.

[Please insert Table 1 about here.]

#### 2.3.4.1 Dependent variable

The dependent variable in all our Model specifications is the natural logarithm of Tobin's Q ratio which is the market value of firm $i$ in year $t$ to its replacement cost:

$$Tobin's\ Q = Q = \frac{Total\_assets - Book + Market\_Value}{Total\_assets} \tag{1}$$

where *Book* is the book value of equity and *Market_Value* is the Market Capitalization. The meaning we ascribe to Tobin's Q is consistent with its interpretation in Hall and Oriani (2006). It indicates the 'market value' of the innovating firm.

#### 2.3.4.2 Explanatory variables: Innovation productivity variables

Our innovation productivity variables are inspired by prior literature (Chan et al., 2001; Deng et al., 1999). We use R&D expense over book value of equity, *RDBE* (worldscope # 05491 is book value per share) (Chan et al., 2001), patents over book value of equity, $Pat/Book$ (Deng et al., 1999) and adjusted patent citation (Gu, 2005) over book value of equity, $Cit/Book$, as our innovation productivity variables.

RDBE is defined as the ratio of the R&D expense of firm $i$ in year $t$ scaled by the book value of equity in year $t$

$$RDBE_{i,t} = \frac{R\&D_{i,t}}{Book_{i,t}} \tag{2}$$

Similarly, we define Pat/Book as the ratio of firm i's patents published in year t scaled by the

book value of equity

$$\frac{Pat_{i,t}}{Book_{i,t}} = \frac{Patents_{i,t}}{Book_{i,t}} \tag{3}$$

In constructing our citation productivity variable, we ensure that the citations count is observable to investors in the market when they make investment decisions. Following Gu (2005), we use citations received in the year $t$ with respect to patents granted in the previous five years. $C_{ik}^{t-j}$ is the number of citations received in year $t$ by patent $k$ for firm $i$ which is granted in year $t-j$ (j=1...5). This number is scaled by the average number of citations received in year $t$ by all patents of the same subcategory granted in year $t-j$ (j=1...5).[14] $N_{t-j}$ is the total number of patents granted in year $t-j$ to firm $i$. This method for adjusting citations propensity to differences in technology fields, grant year and the year in which the citation occurs is in line with Gu (2005) and Hirshleifer et al. (2013). We define Cit/Book as follows:

$$\frac{\text{Cit·i,t}}{\text{Book}_{i,t}} = \frac{\sum_{j=1}^{T}\sum_{k=1}^{N_{t-j}} C_{ik}^{t-j}}{Book_{i,t}} \tag{4}$$

We further dissaggregate our patent and citation productivity variables as 'clean', 'dirty' and 'other'. For example 'clean' patent productivity is defined as follows:

$$\frac{Pat\_clean_{i,t}}{Book_{i,t}} = \frac{Clean\ Patents_{i,t}}{Book_{i,t}} \tag{5}$$

where $Clean\ Patents_{i,t}$ denote the number of clean patents of firm $i$ published in year $t$.

#### 2.3.4.3 Explanatory variables: Innovation efficiency variables

We do not wish to focus exclusively on clean or dirty innovation productivity variables, but also on the efficiency with which research and development (*R&D*) expenditure is used to generate that output. We use two proxies for the measurement of clean/dirty innovation efficiency which are tailored variants on those proxies used in Hirshleifer et al. (2013). First, we study clean/dirty patents scaled by R&D capital, $Pat\_clean/RDC$ and $Pat\_dirty/RDC$.[15] Second, we study adjusted clean/dirty patent citations scaled by R&D expenses, $Cit\_clean/RD$ and $Cit\_dirty/RD$. Hence, whereas Hirshleifer et al. (2013) study innovation efficiency, we focus on clean and dirty innovation efficiency.

$Pat\_clean/RDC$ is defined as the ratio of firm $i$'s clean patents published in year $t$, scaled by its R&D capital in year $t-2$. It can be defined as:

[14] Patent subcategories are defined based on the International Patent Classification.

[15] Research and development expense represents all direct and indirect costs related to the creation and development of new processes, techniques, applications and products with commercial possibilities; Worldscope # 01201.

$$\frac{\text{Pat\_clean˙i,t}}{\text{RDC}_{i,t-2}} = \frac{\text{Clean Patents˙i,t}}{R\&D_{i,t-2} + 0.8 * R\&D_{i,t-3} + 0.6 * R\&D_{i,t-4} + 0.4 * R\&D_{i,t-5} + 0.2 * R\&D_{i,t-6}} \tag{6}$$

The R&D capital is the five year cumulative R&D expenses assuming an annual linear depreciation rate (Chan et al., 2001; Lev et al., 2005). In line with Lev and Sougiannis (1996), we assume a 5 year technology cycle with respect to the benefits of R&D.[16] The time lag between the innovation input (R&D capital) and output (patents) is to account for the average two year application to publication lag documented with respect to US patents (Hall et al., 2001). The use of cumulative R&D expenses in this innovation efficiency measurement is informed by R&D expenses over the preceding five years contributing to successful patent applications in t-2.

As the number of citations made to a firm's clean/dirty patents can reflect the patents' technological or economic importance, we also follow Hirshleifer et al. (2013) to define a new variable which is adjusted clean/dirty patent citations scaled by R&D expenses, $Cit\_clean/RD$ and $Cit\_dirty/RD$. Specifically, $Cit\_clean/RD$ is defined as

$$\frac{\text{Cit\_clean˙i,t}}{\text{RD}_{i,t}} = \frac{\sum_{j=1}^{T} \sum_{k=1}^{N_{t-j}} C_{ik}^{t-j}}{(R\&D_{i,t-2} + R\&D_{i,t-3} + R\&D_{i,t-4} + R\&D_{i,t-5} + R\&D_{i,t-6})} \tag{7}$$

$C_{ik}^{t-j}$ is defined above. The denominator, RD, is the summation of R&D expenses in years $t-2$ to $t-6$. This denominator is informed by the assumption that there is a 2-year application-publication time lag and that only R&D expenditure up to year $t-2$ contributes to patent applications which are published in year $t$.

$Pat\_dirty/RDC$ and $Cit\_dirty/RD$ are defined similarly, focusing on dirty patents only.

#### 2.3.4.4 Control variables: Firm traits

The adopted set of control variables comprises firm traits that can play a role in the market's accordance of stock price value. The set of firm trait variables includes the inverse of book equity, $1/BE$, capital expenditure (Worldscope # 04601) to market value, $CEME$ and advertisement expenditure to market value, $Advert$ (Worldscope # 01101). We control for capital expenditure and advertising expenditure because they are found to explain firm operating performance (e.g., Lev and Sougiannis (1996); Pandit et al. (2011)). The set of firm trait variables also includes *abnormal* earnings, $Earning_{abnormal}$ (the earnings, $E$ is defined as earnings before interest tax depreciation and amortisation, Worldscope # 18198). To obtain abnormal earnings, $Earning_{abnormal}$, earnings, $E$, is adjusted by the corporate income tax rate, $\tau_{i,t}$ (Worldscope #

[16] We set missing R&D to zero throughout but when we repeat our tests with variables with no missing R&D observations we obtain similar findings.

08346) on firm earnings and the annualised risk free rate, $r_t$ (Datastream annualised 90/91 day annualised Treasury bill rate), multiplied by the book value of equity (Ohlson, 1995).

We also include the tax shelter associated with R&D expenditure, *taxRDBE*, as a control variable (Hirshleifer et al., 2013) and substantial R&D growth, *RDG*, (Eberhart et al., 2004). An episode of R&D growth (*RDG*) is captured in a dummy variable which is equal to one if there is an episode of growth of at least 5% in R&D expenditure and a growth of at least 5% in R&D expenditure scaled by total assets relative to the prior year) and is zero otherwise. Eberhart et al. (2004) report significantly positive abnormal stock returns following substantial R&D expenditure growth. Finally, we include time and industry fixed effects in all our regression specifications. We have employed the 48 Fama-French industry classification codes to generate industry dummies. The codes were obtained from Kenneth R. French's website.[17]

### 2.3.5 Descriptive Statistics: Growth in clean and dirty innovation globally

The global rate of growth of production of environmentally friendly 'clean' technologies, vis-a-vis 'dirty' technologies, can be observed in Figure 1, which compares the aggregate clean and dirty patents (and citations made to such patents) published by the US Patent office.[18] This Figure reports a slight increase in the number of dirty patents published during the period 1995-2002, though there is no substantial change in the number of patents published yearly from 2002 to 2012. In contrast, there is a considerable increase in the number of clean patents published with an average growth of 13.58% per year. Figure 2 identifies the top 12 country leaders in clean and dirty innovation.[19] These countries are ranked based on the number of clean and dirty patents published by the US Patent office. All the dirty technology producing countries, except Italy, are also among the top clean technology producing countries. So, if there is a high level of innovation both dirty and clean innovation tend to prevail. A comparison of the aggregate clean and dirty patents published in these countries underscores the rising importance of environmentally friendly technologies in these nations.

[Please insert Figure 1 and Figure 2 about here.]

To assess whether firms have a net incentive or disincentive to produce clean technologies, we construct our innovation productivity (*RDBE*, $Pat/Book$ and $Cit/Book$) and innovation efficiency variables ($Pat/RDC$ and $Cit/RD$) and further disaggregate these variables into 'clean', 'dirty' and 'other' components for investigating their distinct influences on the Tobin's Q of the

[17] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data˙library.html

[18] Clean technologies encompass a markedly larger number of categories than dirty technologies, in our sample. Internet Appendix A, Table A1 reports the list of clean technology categories sampled and Table A2 reports the list of categories for dirty technologies.

[19] The top 12 clean innovation producing countries in descending order are: Japan, USA, Korea, Germany, Taiwan, France, Denmark, Netherlands, Canada, Sweden, Finland and Great Britain. The top 12 dirty innovation producing countries in descending order are: Japan, USA, Germany, Korea, France, Sweden, Finland, Italy, Taiwan, Great Britain, Canada, Netherlands.

firm. The descriptive statistics for these variables are shown in Table 2.

[Please insert Table 2 about here.]

For our dataset, firms on average allocate 4% of their book value of equity to R&D investments. Also, the clean and dirty innovation relative to book value of equity and R&D is a small fraction of total innovation. For instance, while clean and dirty patents over book value of equity account for 3.74% and 0.49%, these same patents over R&D Capital account for 2.62% and 0.68% respectively.

## 2.4 Econometric methodology

In this section, we describe the principal methodologies adopted to elicit the capital market evaluation of clean and dirty innovation. In particular, we describe the extension of the Hall et al. (2005) firm's intangible stock of knowledge function, to account for dis-aggregated clean and dirty innovation productivity and efficiency measures. We also describe Ohlson's accounting based asset valuation model (Ohlson, 1989,9), which serves to inform our Fama-Macbeth two stage (Fama and MacBeth, 1973) estimator work in the robustness tests.

### 2.4.1 Estimation of the Firm-level Market-value stock of knowledge function including Innovation productivity and efficiency variables

We follow Hall et al. (2005) and adopt the firm-level market-value model to evaluate the relationship between R&D investment and the market value of the firm. The chief novelty in our approach consists in the way we apply the model to assess if the stock market recognizes the value of innovation productivity and efficiency in the production of 'clean' and 'dirty' technologies. The market-value model used in Hall et al. (2005), Hall and Oriani (2006) and many other studies on valuation of R&D investments assumes that a firm is valued as a combination of both tangible and intangible assets by the stock market. However, the intangible assets that are created by the R&D investments are often not factored in the computation of the dependent variable, Tobin's Q. The model represents the market value, V, of the firm $i$ at a time $t$ as a function of book value of tangible assets, $A_{i,t}$, replacement value of firm's knowledge assets, $K_{i,t}$, and the replacement value of the other intangible assets, $I^{j}_{i,t}$ and can be represented as below.

$$V_{i,t} = V(A_{i,t}, K_{i,t}, I^{1}_{i,t}, \ldots, I^{n}_{i,t}) \tag{8}$$

Assuming assets can be written in an additive and linearly separable fashion and neglecting the other intangible assets, the market-value model is expressed as

$$V_{i,t} = b(A_{i,t} + \gamma K_{i,t})^{\sigma} \tag{9}$$

where $\sigma$ accounts for the non-constant scale effects in the market-value function, $\gamma$ represents the shadow value of knowledge assets relative to a firm's tangible assets and $b$ denotes the average market valuation coefficient of total assets of a firm and can be interpreted to account for a firm's monopoly position and its differential risk (Grandi et al., 2009). Simplifying the representation of the model by taking the natural logarithm on both sides of the equation and assuming that $\sigma$=1 we get the following model

$$logV_{i,t} = logb + log(A_{i,t}) + log(1 + \gamma\frac{K_{i,t}}{A_{i,t}}) \tag{10}$$

which further simplifies to

$$logQ_{i,t} = log(\frac{V_{i,t}}{A_{i,t}}) = logb + log(1 + \gamma\frac{K_{i,t}}{A_{i,t}}) \tag{11}$$

where $Q_{i,t}$ stands for Tobin's Q. From the above model, one can estimate the average effect of a unit currency invested in knowledge assets on the firm's market value.

In creating our innovation productivity and efficiency variables, we consider that the full value of R&D investments can be captured from investment in R&D to creation of patents to efficiency of R&D investment in generating patents, to the generation of citation and finally the efficiency of R&D investment in creating citations. So, in our specifications we use R&D over book value of equity $(RDBE)$ as a proxy for R&D productivity; patents over book value of equity $(Pat/Book)$ and patents over R&D Capital $(Pat/RDC)$ as proxies for patent productivity and efficiency; and citations over book value of equity $(Cit/Book)$ and citations over RD $(Cit/RD)$ as proxies for citation productivity and efficiency. We further disaggregate these variables into 'clean', 'dirty' and 'other' components to determine their relative importance in assessing the market value of the firm.

We first assess the impact of each individual innovation productivity and efficiency variable on the Tobin's Q of the firm by estimating various specifications derived from the Models

$$\begin{aligned} logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \\ \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it} \end{aligned} \tag{12}$$

and

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it} \quad (13)$$

Year and industry dummies represent time and industry fixed effects. We dis-aggregate the main innovation variables into 'clean', 'dirty' and 'other' components and examine whether the stock market attaches any importance to these technology classes separately. We also analyze the relative importance of each of the innovation productivity and efficiency variable. For this, we estimate various specifications of the following Models:

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it} \quad (14)$$

and

$$logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it} \quad (15)$$

where $Pat^*$ and $Cit^*$ denote the 'clean', 'dirty' or 'other' knowledge asset.

### 2.4.2 Estimation of Market value as a function of Innovation productivity and efficiency stocks using Ohlson's accounting based asset valuation Model

We adapt the Ohlson (1989) accounting-based asset valuation model to examine whether, and, if so, to what extent, the stock market assimilates the information content in clean and dirty innovation production and efficiency.[20] This model allows a test of whether clean and dirty innovation expenses explain market value and of any difference between their market value contributions. Ohlson (1989) derives the following valuation equation:

$$M_{i,t} = BE_{i,t} + \beta_0[E_{i,t}(1 - \tau_{i,t}) - r * BE_{i,t}] + \beta_1[\tau_{i,t} RD_{i,t}] + \alpha * Z_{i,t} \quad (16)$$

[20]This general asset pricing framework is also used in Barth et al. (1998); Sougiannis (1994); Ohlson (1995) and Hirshleifer et al. (2013) among others. It is recommended in Brennan's 1991 review paper (Brennan, 1991).

where $M_{i,t}$ is the market value of the $i^{th}$ firm at time t. $[E_{i,t}(1-\tau_{i,t})-r*BE_{i,t}]$ is a measure of abnormal earnings discussed above and initially defined in Ohlson (1989); $[\tau_{i,t}RD_{i,t}]$ accounts for the tax shelter associated with *R&D* expenditure; $Z_{i,t}$ is a vector of other information variables. Other variables are as defined above.

In our adaptation of this accounting-based asset valuation model, we use the natural logarithm of Tobin's Q as the dependent variable and we include 'clean', 'dirty' and 'other' innovation productivity and efficiency variables, and the control variables used in Hirshleifer et al. (2013) as our vector of controls (RDG, Earning˙abnormal, invBE, CEME, Adverts, taxRDBE[21]).

We run non-linear least squares regressions in line with Hall et al. (2005), see equations 17 and 18, as well as Fama-MacBeth (1973) annual cross-sectional regressions at the firm level, see equations 19 and 20. We specify and estimate equations 19 and 20 following Hirshleifer et al. (2013), to test if our findings are invariant to an alternative estimator: the Fama-MacBeth (1973) estimator. Our robustness tests regression specifications are derived from the following models:

$$\begin{aligned} logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} \\ + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \\ \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it} \end{aligned} \tag{17}$$

$$\begin{aligned} logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} \\ + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \\ \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it} \end{aligned} \tag{18}$$

$$\begin{aligned} logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} \\ + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \\ \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it} \end{aligned} \tag{19}$$

[21] See Table 3 of the definition of these variables.

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it} \quad (20)$$

$Pat^*$, and $Cit^*$, are 'clean', 'dirty' or 'other' patents and citations.

## 2.5 Empirical findings

This section presents our baseline empirical results. It then presents results of robustness tests on a variety of dimensions: alternative estimators, sub-samples of firms which have conducted both clean and dirty innovation, accounting for firm traits and emerging technology innovation and tests for whether comparable findings hold for European patents. We discuss the baseline results in subsection 2.5.1. The results of the robustness tests are discussed in subsections 2.5.2 to 2.5.7.

### 2.5.1 Baseline regressions: Association between Tobin's Q and Innovation productivity and efficiency variables

Tables 3 and 4 report the results for the non-linear regression specifications which are derived from the firm-level market value model and are similar to those reported in Hall et al. (2005). We first determine the innovation productivity and efficiency variables' association with a firm's Tobin's Q (Table 3), and, then, disaggregate these variables into clean, dirty and other components to assess their distinctive associations with a firm's Tobin's Q (Table 4). All our model specifications include time and industry fixed effects. Since R&D productivity is highly correlated with the firm's individual effect, we exclude firm fixed effects to sidestep over-correction (Hall et al., 2005).

Table 3 reports the results for specifications derived from equations (12) and (13). The results suggest that, on average, R&D, patent and citation productivity (*RDBE*, *Pat*/*Book* and *Cit*/*Book*) positively correlate to Tobin's Q.[22] In the light of the new international data examined, this corroborates the main findings reported in Hall et al. (2005). We also assess the association between the efficiency of R&D investments in generating patents and citations with the Tobin's Q (Hirshleifer et al., 2013) to find that innovation efficiency variables (*Pat*/*RDC*, *Cit*/*RD*) are also positively associated with Tobin's Q. To determine the association of these variables with the Tobin's Q, we estimate the corresponding semi-elasticities, the results of

[22]Please refer to Table C1 in the Internet Appendix C which reports consistent findings for European patents, and Table D1 of the Internet Appendix D which shows consistent results from a Fama-Macbeth regression framework.

which can be found in Table B1 in the Internet Appendix B. For example, the semi-elasticities with respect to citation over book ($Cit/Book$) for specification 3 suggest that an additional citation per million dollars of book value of equity is associated with an increment of 1.1% ($e^{.1073}$) in Tobin's Q, respectively. Similarly, for specification 4 and 5, we find that the patents over R&D capital ($Pat/RDC$) and citations over RD ($Cit/RD$) are positively associated with the Tobin's Q with an economic relation of approximately 1% ($e^{.0030}, e^{.0109}$).[23]

[Please insert Table 3 about here.]

To determine whether the capital markets incentivize clean innovation vis-a-vis dirty innovation, we disaggregate patents over book ($Pat/Book$), citations over book ($Cit/Book$), patents over R&D capital ($Pat/RDC$), and citations over RD ($Cit/RD$) into clean, dirty and other components. We estimate the semi-elasticities for each specification reported in Table 4 with respect to the dis-aggregated innovation and innovation efficiency variables to determine their association with the Tobin's Q. For the first specification reported in Table 4, we find that the clean patents over book ($Pat\_clean/Book$) is positively associated with the Tobin's Q at an economic value of 3.77% ($e^{1.3270}$). We also find that the clean citation over book ($Cit\_clean/Book$) is positively associated with Tobin's Q at an economic value of 1.27% (specification 2 of Table 4). Additionally, we disaggregate our innovation efficiency variables and find that the clean citations over RD ($Cit\_clean/RD$) is positively related to the dependent variable with an economic value of 1% (specification 4 of Table 4). We find that the clean patents over R&D capital ($Pat\_clean/RDC$) is positively related to Tobin's Q, though this result is not statistically significant (specification 3 of Table 4). However, efficiency of R&D investments in generating dirty patents decreases the market value of the firm to the tune of 0.97% economic value (specification 3 of Table 4).[24] Significantly, the t-test for the difference between coefficients of clean and dirty patents over book ($Pat\_clean/Book - Pat\_dirty/Book = 0$), patents over R&D capital ($Pat\_clean/RDC - Pat\_dirty/RDC = 0$), citations over book ($Cit\_clean/Book - Cit\_dirty/Book = 0$), and citations over RD ($Cit\_clean/RD - Cit\_dirty/RD = 0$) are all statistically different from zero at a 5% level. The results for semi-elasticities for Table 4 are consistent and t-tests can be found in Tables B2 and B6 (Panel A) in the Internet Appendix B.[25] [26]

[23] Please refer to Table B1 in the Internet Appendix B.

[24] Please refer to Table C2 in the Internet Appendix C which reports consistent findings for European patents and Table D2 of the Internet Appendix D which shows consistent results from a Fama-Macbeth regression framework.

[25] Tables E1 and E2 of Internet Appendix E, using a non-linear least squares estimator and a Fama-Macbeth regression specification, report consistent results with future operating profit, i.e., earnings before interest, taxes, depreciation, and amortization (EBITDA), as a response variable.

[26] Tables K1 and K2 of Internet Appendix K, using a non-linear least squares estimator and a Fama-Macbeth regression specification, report consistent results with adjusted patent citations measeures as the key innovation measures. The analysis in Internet Appendix K employs patent data set from the US patent office which covers 2526 US listed firms during 1995 to 2012. While previous empirical studies in largely using patent counts to indicate innovation have ignored the differences across industries, the present analysis accounts for this concern by employing econometric models with adjusted patent citations and industry fixed effects. It is widely accepted that citations of a firm's patents indicate the technological and economic significance of the innovation (Harhoff

[Please insert Table 4 about here.]

### 2.5.2 Do the main results hold using a Fama-Macbeth two-step estimator?

As an alternative econometric approach to the firm-level market value model used in Hall et al. (2005) and other studies on valuation of R&D investments, we adopt the popular Fama-MacBeth estimator (Fama and MacBeth, 1973) to assess the Models in Table 4 and this confirms the prevalence of a clean innovation premium. The economic upshot of clean innovation productivity and efficiency is similar to that reported in Table 4, with the exception of clean patent productivity ($Pat\_clean/Book$), which is three times higher than the corresponding clean patent productivity ($Pat\_clean/Book$) association reported in Table 4[27].

[Please insert Table 5 about here.]

### 2.5.3 Do the main results hold using a sub-sample of firms which produces both clean and dirty technologies?

A potential issue is that in the sector of electricity generation, dirty firms tend to be large incumbents while clean firms are typically smaller entrants. In the absence of firm fixed effects, the results could therefore be driven by unobserved intrinsic and time-invariant differences in the type of firms conducting clean or dirty innovation which are not controlled for in the regressions. Therefore, we estimate the models reported in Table 4 for the sub-sample of firms producing both clean and dirty technologies. This allows us to assess if there is clean innovation premium *within* firms producing both clean and dirty technologies.

The results are reported in Table 6. We find that our results are robust with respect to clean patent ($Pat\_clean/Book$) and citation ($Cit\_clean/Book$) productivity variables, respectively. We also find that the efficiency of R&D investments in generating dirty patents ($Pat\_dirty/RDC$) and citations ($Cit\_dirty/RD$) decrease the Tobin's Q of the firm to the tune of 0.98%.[28] Further, the difference between coefficients of clean and dirty patent ($Pat\_clean/Book - Pat\_dirty/Book$) and citation productivity ($Cit\_clean/Book - Cit\_dirty/Book$) and the difference between clean and dirty coefficients of citation efficiency ($Cit\_clean/RD - Cit\_dirty/RD$) variables are posi-

et al., 2003; Hall et al., 2005; Hirshleifer et al., 2013). Therefore, Tables K1 and K2 employ adjusted patent citations as proxy for the firm's technology.

[27]The first specification of Table 5 suggests that a unit increase in Pat_clean/Book is associated with an increase of 2.319 in the natural logarithm of Tobin's Q (log Q). So, a one unit increase in Pat_clean/Book is associated with an increase of 10.17% ($e^{2.319}$) in Tobin's Q. Since the non-linear estimation of the corresponding model (first specification of Table 4) suggests that Pat_clean/Book is positively associated with Tobin's Q with an economic impact of 3.77%, we infer that the economic impact derived from Table 5 is approximately three-fold of the corresponding Pat_clean/Book derived from Table 4.

[28]We estimate our baseline Models for the sample of firms with non-zero patents (See Tables F2 and F3 in the Internet Appendix F). We find a positive and statistically significant association between innovation productivity and efficiency variables with the Tobin's Q of the firm and further, find a positive and significant association between clean innovation productivity ($Pat\_clean/Book, Cit\_clean/Book$) variables and clean citation efficiency ($Cit\_clean/RD$) variables with the Tobin's Q of the firm, respectively.

tive and statistically different from zero at a the 5% level. Also, the difference in the premia associated with the efficiency with which R&D investments generate clean and dirty patents ($Pat\_clean/RDC - Pat\_dirty/RDC$) is statistically different from zero at 10%. The results for semi-elasticities for Table 6 and related t-tests can be found in Tables B3 and B6 (Panel B) in the Internet Appendix B.

We conclude from this test that the result is not simply driven by unobserved heterogeneity between firms conducting clean or dirty innovation, but that a clean innovation premium holds *within* diversified firms conducting both types of innovation.

[Please insert Table 6 about here.]

### 2.5.4 Do the main results hold explicitly accounting for emerging technologies in our regressions?

We are concerned that the estimates of clean innovation productivity and efficiency may be relaying the effect of emerging technologies more generally on the firm's Tobin's Q.[29] [30] Emerging technologies are new and disruptive innovations such as Information technologies, robots or nanotechnologies, that are likely positively associated with both the firm's Tobin's Q as well as with clean technologies, if some firms specialize in emerging technologies in general, which encompass clean technologies. Hence, the omission of emerging technologies may upwardly bias the estimates of clean innovation productivity and innovation efficiency. The patent classification codes used to extract emerging patents from the database is presented in Table A3 in the Internet Appendix A.

Therefore, we disaggregate the 'other patents' into 'emerging' and 'mature' technologies,[31] and we extend the Models reported in Table 4 to include the patent and citation productivity ($Pat\_emtech/Book$, $Cit\_emtech/Book$) in emerging technologies and the corresponding efficiency variables ($Pat\_emtech/RDC$, $Cit\_emtech/RD$) as controls. Table 7 reports the findings. We find no substantial change in the estimates of clean innovation productivity and innovation

[29] We thank a reviewer for highlighting that due to a life-cycle and a decreasing returns channel at the patent level, mature technologies (e.g. dirty innovation) can experience decreasing returns, and a weaker Tobin's Q association than clean innovation. This can potentially account for our main finding of a clean innovation premium. Table G1, of the Internet Appendix G, reports that for emergent technologies, presumably in the early phase of their life cycle, there is no clean innovation premium for patents (Column 1) but that a clean innovation premium is still evident for clean innovation citations (Column 2). This suggests some evidence in support of a patent technology category life-cycle mechanism to account for a clean innovation premium. Note that the paper is focused on establishing whether there is a clean innovation premium and does not claim to establish the drivers of such a premium - see the discussion in the concluding section.

[30] Tables J1 and J2 in the Internet Appendix J include innovation measures with respect to grey technologies. Grey technologies make dirty innovation "less dirty" (e.g., making a combustion engine more efficient). Grey technologies show a positive relation with Tobin's Q and their inclusion does not compromise the main result of a clean innovation premium.

[31] For the sake of simplicity we denote 'mature' technologies as 'other' technologies when we include innovation productivity and efficiency variables with respect to emerging technologies in our Models.

efficiency. This substantiates the results reported in Table 4.[32] We also find that the t-test for the difference between coefficients of clean and dirty patents over book ($Pat\_clean/Book - Pat\_dirty/Book = 0$), patents over R&D capital ($Pat\_clean/RDC - Pat\_dirty/RDC = 0$), citations over book ($Cit\_clean/Book - Cit\_dirty/Book = 0$), and citations over RD ($Cit\_clean/RD - Cit\_dirty/RD = 0$) are statistically different from zero at a 5% level. The results for semi-elasticities for Tables 7 and related t-tests can be found in Tables B4 and B6 (Panel C) in the Internet Appendix B.[33]

[Please insert Table 7 about here.]

### 2.5.5 Do the main results hold explicitly accounting for accounting-based asset valuation firm-level traits in our regressions?

As a further robustness test to deal with a potential omitted variable bias in the absence of firm fixed effects, we extend the non-linear regression models reported in Table 4 by including firm traits in line with the Ohlson's accounting based asset valuation model cited in Hirshleifer et al. (2013). In this heavily parameterized setting, our main results hold well in respect to clean and dirty citations over RD ($Cit\_clean/RD$, $Cit\_dirty/RD$), as indicated in specification 4 of Table 8. We also include patent and citation productivity and efficiency with respect to emerging technologies (Specifications 5-8 of Table 8), and again find that our results are robust with respect to clean and dirty citations over RD ($Cit\_clean/RD$, $Cit\_dirty/RD$), as indicated in specification 8 of Table 8. The estimates of clean citation efficiency, $Cit\_clean/RD$, reported in specifications 4 and 8 of Table 8 are similar to the one reported in Table 4 having the same economic association of 1.04% with a firm's Tobin's Q. We also find that the efficiency of R&D investments in generating dirty citations ($Cit\_dirty/RD$) decreases the Tobin's Q of the firm to the tune of 0.99%. For specifications 4 and 8 we find that difference between coefficients of clean and dirty citations over RD ($Cit\_clean/RD - Cit\_dirty/RD = 0$) are statistically different from zero at a 5% level. The results for semi-elasticities for Table 8 and related t-tests can be found in Tables B5 and B6 (Panel D and E) in the Internet Appendix B.

[Please insert Table 8 about here.]

Further, these Models are also estimated using the Fama-MacBeth estimator and our main result that the stock market accords significantly more value to clean as opposed to dirty innovation

[32] We estimate the Models reported in Table 7 for the sample of firms producing both clean and dirty patents and find that our result of clean innovation premium holds with respect to patent and citation productivity (See Table F5 in the Internet Appendix F). We also estimate these Models for the sub-sample of firms with non-zero patents and find clean innovation premium with respect to innovation productivity variables and citation efficiency variables (See Table F4 in the Internet Appendix F).

[33] We also adopt the Fama-MacBeth estimator to assess these Models. We find that the economic value of clean innovation productivity and efficiency is similar to those derived from Table 7. Please refer Table D2 in the Internet Appendix D.

productivity and innovation efficiency remain unchanged.[34] [35]

As demand in the market and generic government policies inform a firm's decision to innovate in a particular area, we posit that the 5-year change in the Environmental policy stringency score (Botta and Koźluk, 2014) would proxy for the appetite, for clean innovation, of the investors and consumers. Therefore, we add the difference between one-year and six-year lag of Environmental policy stringency score of the US ($EPSlag1 - EPSlag6$) and emerging technology variants of innovation productivity and efficiency variables to the baseline regression models (Models in Table 4) and find that there is still a clean innovation premium with respect to efficiency of R&D investment in generating citations. We argue that this finding is economically relevant as citations show the importance of a particular innovation and further propel innovation in that area.[36]

### 2.5.6 Do the main results hold explicitly accounting for a managerial selection bias?

In our study, sample selection bias may arise if managers choose to innovate in clean technologies more relative to dirty technologies. Therefore, to address sample selection we adopt the Heckman two stage 1979 regression approach (Heckman, 1979). Table 9 reports the related findings. In the first stage, we model the likelihood of a firm to conduct clean innovation using a Probit model. The dependent variable for the first stage is $Clean\_firm$, which is a dummy variable that takes the value 1 if a firm has a clean patent published by the USPTO during the period 1995-2012 and 0 otherwise. We regress $Clean\_firm$ on $Emtech\_firm$,[37] $Total\_assets$, $EPSlag1 - EPSlag6$, the full set of control variables, year and industry dummies:

$$Clean\_firm = \alpha + \gamma_1 Emtech\_firm_i + \gamma_2 RDG_{it} + \gamma_3 invBE_{it} + \gamma_4 taxRDBE_{it} + \gamma_5 CEME_{it} \quad (21)$$
$$+\gamma_6 Earning_{abnormal_{it}} + \gamma_7 Adverts_{it} + \gamma_8 Total\_assets + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$$

For the second stage we use Models 1 and 2 of Tables 4 and 8 and include the inverse Mills ratio (bias correction term), obtained from the first stage, as an explanatory variable. We find that our inference of a clean innovation premium remains, despite this correction.

[Please insert Table 9 about here.]

[34] Please refer to Table D3 in the Internet Appendix D.

[35] The main results hold even when we construct the innovation productivity and efficiency variables with respect to the grant date instead of publication date.

[36] Please refer Table F1 in the Internet Appendix F.

[37] Emtech_firm is a dummy variable that takes the value 1 if a firm has an emerging technology patent published by the USPTO during the period 1995-2012 and 0 otherwise.

### 2.5.7 Do the main results hold for European patents?

To check if our results hold in a different jurisdiction, we run the robustness tests for the patents and citations published by the European Patent Office (EPO). We find a positive association between clean patent productivity ($Pat\_clean/Book$) and Tobin's Q and this result is statistically significant at 5%. We also find a negative and significant association between dirty citation productivity ($Cit\_dirty/Book$) and efficiency variables ($Cit\_dirty/RD$) with the Tobin's Q.[38]

Following the test presented in section 5.4, we also include the patent and citation productivity and efficiency with respect to emerging technologies and find that the results do not change substantially.[39] These Models were estimated using a non-linear least squares estimation method.

Additionally, we estimate these Models using a Fama-MacBeth estimator and find that our main results hold with regard to clean and dirty patent productivity and efficiency. We also extend these models to include a patent and citation productivity and efficiency with respect to emerging technologies ($Pat\_emtech/Book$, $Cit\_emtech/Book$, $Pat\_emtech/RDC$, $Cit\_emtech/RD$) and thus find that there is a positive and significant association between generating clean relative to dirty patents efficiently and Tobin's Q.[40]

Further, we estimate the association of clean and dirty innovation productivity and efficiency variables with Tobin's Q of the firm, while controlling for emerging technology variants of innovation productivity and efficiency variables and firm traits in line with the Ohlson's accounting based asset pricing model cited in Hirshleifer et al. (2013). In this heavily parameterized setting, our main results hold well in respect to clean and dirty patent productivity and efficiency ($Pat\_clean/Book$, $Pat\_dirty/Book$, $Pat\_clean/RDC$ and $Pat\_dirty/RDC$), as indicated in specifications 1, 3, 5 and 7 of Table C7 in the Internet Appendix C.[41]

## 2.6 Conclusion and Discussion

Innovation productivity is critically important for firm- and national-level competitiveness in international markets (Porter, 1992). Innovation productivity to curtail, and ultimately reverse, environmental degradation (*i.e.* 'clean' innovation) can prove vital to establish a sustainable market economy around the world (Allen and Yago, 2011; IPCC, 2014). Such a sustainable market economy will mitigate market failures and serve to protect air, water, fisheries, wildlife, and biodiversity. In this paper, we raise the question of whether there is an economic incentive

[38] Please refer Table C2 in the Internet Appendix C.

[39] Please refer Table C3 in the Internet Appendix C.

[40] Please refer Tables C5 and C6 in the Internet Appendix C.

[41] Please refer to Tables H1 to H7 in the Internet Appendix H, which report findings that the main results are invariant to time, industry, firm *and* country level control variables. Please refer to Tables I1 to I6 in the Internet Appendix I, which report findings that the main results are invariant to using book value of assets as opposed to the book value of equity as a denominator. We thank a reviewer and an Associate Editor for suggesting these latter tests of the robustness of our main findings.

for firms to pursue strategies of clean environmentally-supportive innovation, as opposed to carbon-emitting dirty innovation activities.

We use a unique dataset covering 15,217 listed firms across 12 countries to measure the relationship between market value and innovation activity. We disaggregate annual patent counts by technology, distinguishing between clean, dirty and other technologies (including emergent technologies). Our dataset also includes patent citation data which is used to proxy for patent quality.

We start by verifying the value accorded by the capital market to generic innovation and innovation efficiency internationally, in the non-linear regression model setting of Hall et al. (2005). This serves to establish the validity of our data and empirical set-up.

Our main contribution is that we elicit capital market evaluations associated with the disaggregated innovation productivity measures (Deng et al., 1999; Chan et al., 2001) and innovation efficiency measures (Hirshleifer et al., 2013) to account for 'clean' and 'dirty' innovation production and efficiency. We report that 'clean' innovation efficiency is typically associated with an economically important and positive Tobin's Q, while the capital market ascribes no (or a negative) market value influence to 'dirty' innovation efficiency.

The relative Tobin's Q association of 'clean' vis-a-vis 'dirty' innovation is significant and economically important across innovation measurements. These main results are invariant with respect to a range of model specifications, a focus on European as opposed to United States patents, sub-samples of firms which conduct both clean and dirty innovation, estimation strategies, and controlling for firm traits frequently used in respect to asset pricing.

Our question is whether there is a clean innovation premium, consistent with the objective for a long-term de-carbonization of the international economy. We do not, thus, aim to discern, from the data, why a clean or dirty innovation premium can prevail. The question we raise is nonetheless important. Its resolution is also not straightforward. We, with novelty, avail of a compelling litmus test to resolve the raised question: the information content of equity market price signals. As such, we meaningfully address this complex and important question, and report strong and robust evidence of a clean innovation premium.

Several competing or complementary explanations can drive the existence of a clean innovation premium. A first possible explanation is that clean patents signal greater growth opportunities than dirty patents, in a world that is increasingly constrained by climate change mitigation policies. Clean investors might also need to invest more in the future to realise the value of their patent stock than firms producing dirty patents. A major competing candidate, however, is the existence of decreasing marginal returns to R&D, which could contribute to smaller effects of incremental patenting on Tobin's Q over time, as 'dirty' technologies are more mature than 'clean' innovations. It could also be that patents on 'clean' technologies are more difficult to

produce than patents in 'dirty' technologies, which would be rewarded by the market (although this argument goes against the assumption of decreasing marginal returns from R&D efforts).

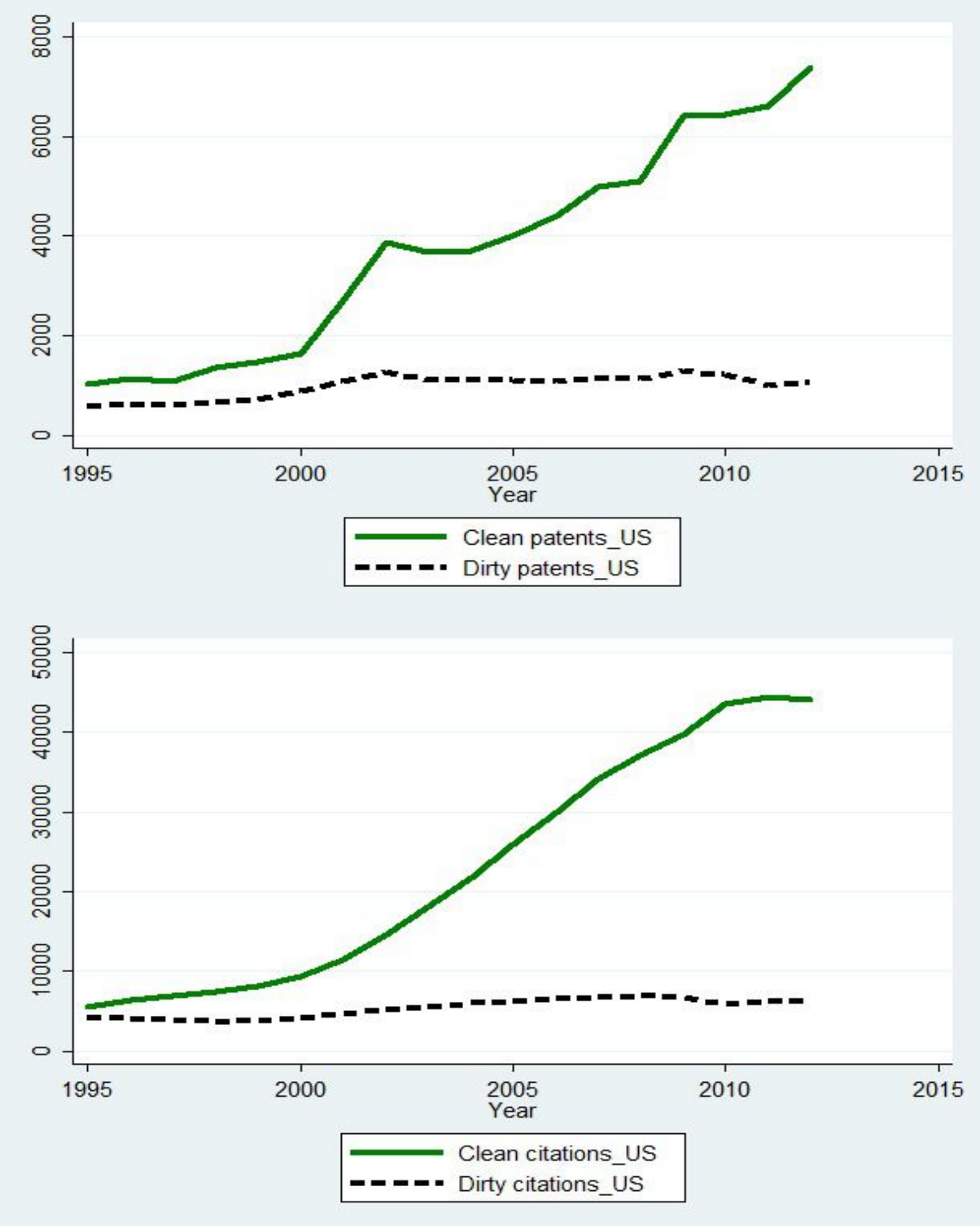
Therefore, an important avenue for research is to empirically investigate the drivers behind the clean innovation premium uncovered in this paper. This is left for future work.

## 2.7 Tables and Figures

## Table 1: Variable Definitions

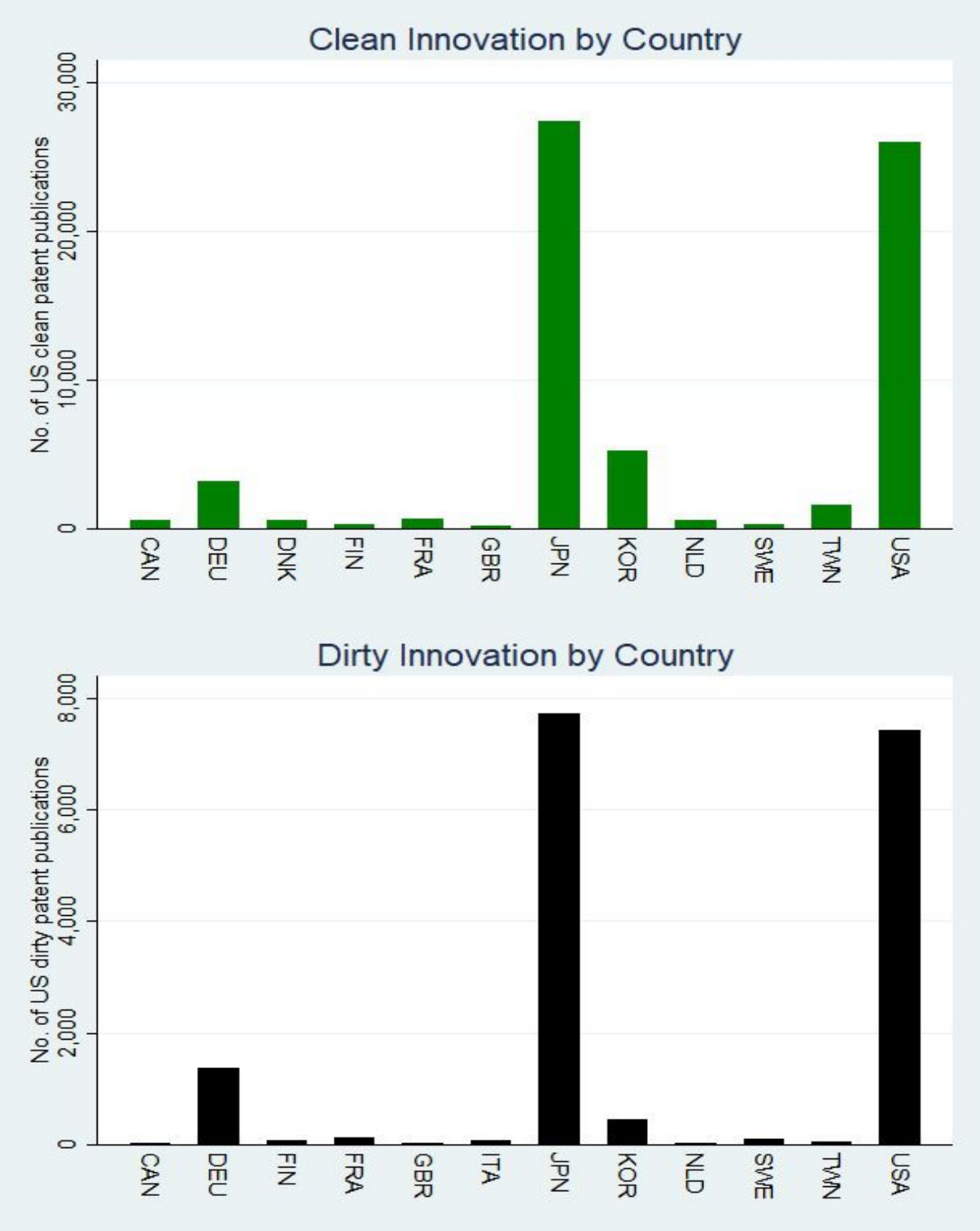
| Variable | Definition |
|---|---|
| Measures of firm value | |
| Tobin's Q | Market value of the firm to the book value of tangible assets $(Total\_assets - Book + Market\_Value)\ /(Total\_assets)$. |
| Total_assets (millions of $) | Total Assets represents the sum of total current assets, long term receivables, investment in unconsolidated subsidiaries, other investments, net property plant and equipment and other assets. |
| Market_Value | Total market value of the company based on year end price and number of shares outstanding converted to U.S. dollars using the year end exchange rate. |
| Book (millions of $) | Book value of equity. |
| Measure of R&D Productivity | |
| RDBE | Research and Development expense divided by Book. |
| Measures of Innovation Productivity | |
| Pat/Book | Number of US patents of the firm, in any patent category, divided by Book. |
| Pat*/Book | As per Pat/Book but US patent category is *: clean, dirty, other or emerging technologies. |
| Cit/Book | The numerator is the number of citations received in year t by US patent k, granted in year t-j (j=1-5) scaled by the average number of citations received in year t by all patents of the same subcategory granted in year t-j (j=1-5). This number is summed over the total number of patents granted in year t-j to firm i. The numerator is divided by the book value of equity. |
| Cit*/Book | As per Cit/Book but US patent category is *: clean, dirty, other or emerging technologies. |
| Measures of Innovation Efficiency | |
| Pat/RDC | Number of US patents of the firm divided by the 5-year cumulative R&D expenses, observed in year t-2, assuming a depreciation rate of 20% per annum. |
| Pat*/RDC | As per Pat/RDC but US patent category is *: clean, dirty, other or emerging technologies. |
| Cit/RD | The numerator is the number of citations received in year t by US patent k, granted in year t-j (j=1-5) scaled by the average number of citations received in year t by all patents of the same subcategory granted in year t-j (j=1-5). This number is summed over the total number of patents granted in year t-j to firm i. The numerator is divided by the summation of R&D expenses in years t-3 to t-7. |
| Cit*/RD | As per Cit/RD but US patent category is *: clean, dirty, other or emerging technologies. |
| Firm traits | |
| invBE | Inverse of Book. |
| CEME | Capital expenditure (funds used to acquire fixed assets other than those associated with acquisitions) to Market Value of Equity. |
| Adverts | Advertising expenditure to Market Value of Equity. |
| RDG | R&D growth; An episode of R&D growth (RDG) is captured in a dummy variable which is equal to one if there is an episode of growth (R&D expenditure is greater than 5% of total assets and of total sales and there is a growth of at least 5% in R&D expenditure and a growth of at least 5% in R&D expenditure scaled by total assets relative to the prior year) and is zero otherwise (Total sales measured in millions of $, is the gross sales and other operating revenue less discounts, returns and allowances). |
| Earning˙abnormal | Abnormal earnings; earnings before interest tax depreciation and amortization, E, is adjusted by the corporate income tax rate, $\tau_{i,t}$ on firm earnings and the annualized risk free rate, $r_t$, multiplied by the book value of equity is deducted. |
| taxRDBE | Tax shelter associated with R&D expenditure |
| Regulation | |
| EPS | Environmental Policy Stringency Index (Botta and Kozluk, 2014); This index takes the value from 0 (least stringent) to 6 (most stringent) and is a country-specific stringency measure. |

**Figure 1: Clean and dirty patents and citations**



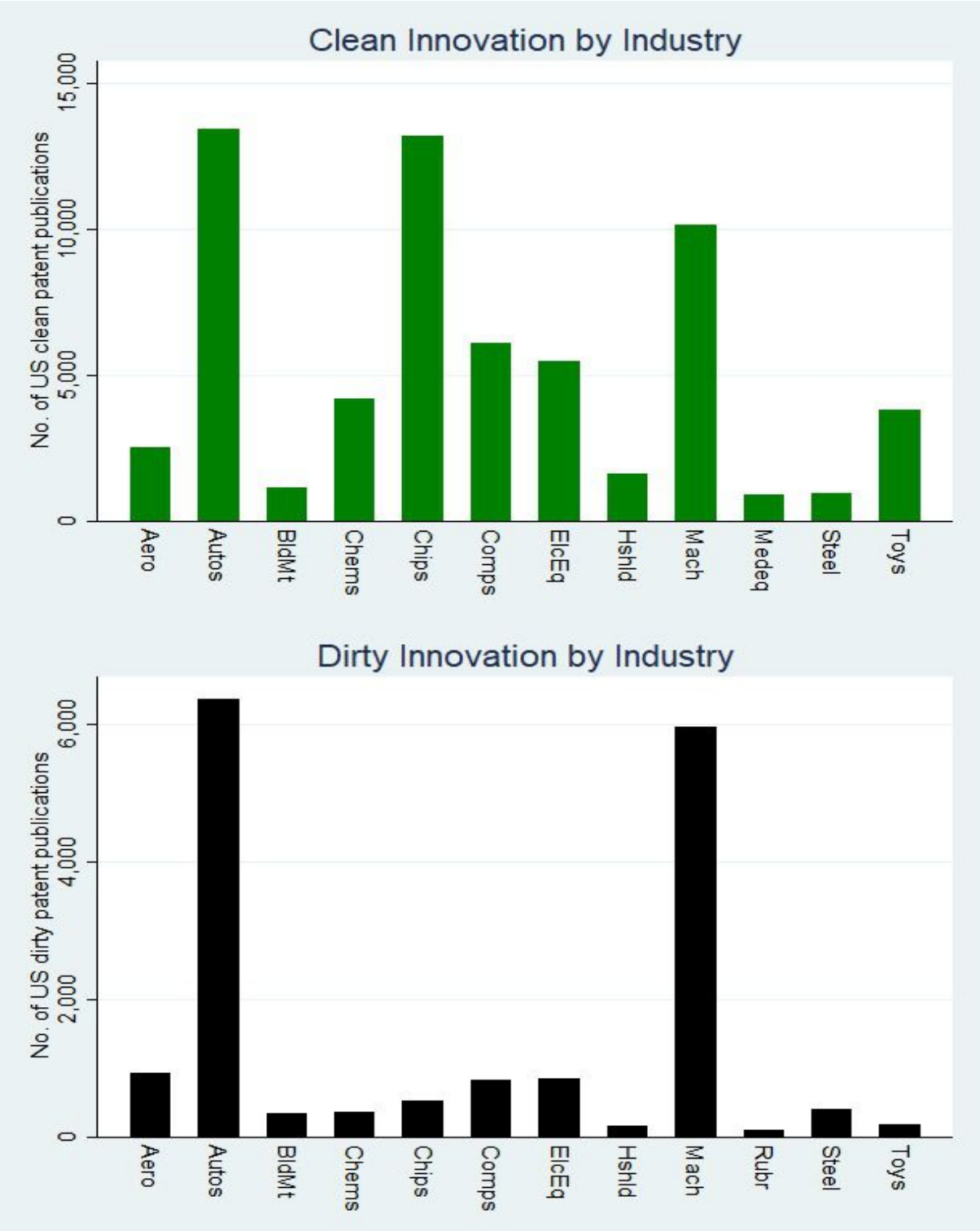

Notes. The Figure shows, over time, the number of published patents in clean and dirty technologies in the US (upper Panel) and shows related citations, accumulated in a 5-year window, in regard to clean and dirty innovations (lower Panel). We refer to Clean (Dirty) patents_US as the total number of clean (dirty) patents published by the USPTO during the period 1995-2012. We refer to Clean (Dirty) citations_US as the number of clean (dirty) patent citations of the firm, related to patents granted in the past 5 years by the USPTO.

**Figure 2: Clean and dirty patent productivity by country**




Notes. The Figure shows the number of published patents in clean and dirty technologies held by 12 leading clean technology producing countries (upper Panel) and 12 leading dirty technology producing countries (lower Panel). The top 12 clean innovation producing countries in descending order are: Japan, USA, Korea, Germany, Taiwan, France, Denmark, Netherlands, Canada, Sweden, Finland and Great Britain. The top 12 dirty innovation producing countries in descending order are: Japan, USA, Germany, Korea, France, Sweden, Finland, Italy, Taiwan, Great Britain, Canada, Netherlands.

**Figure 3: Clean and Dirty Patent productivity by Industry**




Notes. The Figure shows the top 12 leading clean technologies producing industries (upper Panel) and the top 12 leading dirty technologies producing industries (lower Panel) in the 12 leading clean technology producing countries. The top 12 clean innovation producing industries in descending order are: Autos (Automobile), Chips (Electronic equipment), Mach (Machinery), Comps (Computers), ElcEq (Electrical equipment), Chems (Chemicals), Toys (Recreation), Aero (Aircraft), Hshld (Consumer goods), BldMt (Construction materials), Steel (Steel) and Medeq (Medical equipment). The top 12 dirty innovation producing industries in descending order are: Autos (Automobile), Mach (Machinery), Aero (Aircraft), ElcEq (Electrical equipment), Comps (Computers), Chips (Electronic equipment), Steel (Steel), Chems (Chemicals), BldMt (Construction materials), Toys (Recreation), Hshld (Consumer goods) and Rubr (Rubber and Plastic).

**Table 2: Summary Statistics**

| VARIABLES | N | Mean | Standard deviation |
|---|---|---|---|
| Innovation intensity | | | |
| RDBE | 283,254 | 0.0426 | 1.4510 |
| Pat/Book | 186,710 | 0.0267 | 2.2100 |
| Pat_clean/Book | 186,710 | 0.0010 | 0.1870 |
| Pat_dirty/Book | 186,710 | 0.0001 | 0.0060 |
| Pat_emtech/Book | 186,710 | 0.0062 | 0.7180 |
| Pat_other/Book | 186,710 | 0.0256 | 2.0390 |
| Cit/Book | 186,710 | 0.1320 | 7.8290 |
| Cit_clean/Book | 186,710 | 0.0048 | 0.4940 |
| Cit_dirty/Book | 186,710 | 0.0006 | 0.0298 |
| Cit_emtech/Book | 186,710 | 0.0330 | 3.5470 |
| Cit_other/Book | 186,710 | 0.1270 | 7.5040 |
| Innovation efficiency | | | |
| Pat/RDC | 283,253 | 0.0855 | 7.8890 |
| Pat_clean/RDC | 283,254 | 0.0022 | 0.1180 |
| Pat_dirty/RDC | 283,254 | 0.0006 | 0.0595 |
| Pat_emtech/RDC | 283,254 | 0.0073 | 0.2500 |
| Pat_other/RDC | 283,253 | 0.0827 | 7.8860 |
| Cit/RD | 283,254 | 0.2100 | 8.5060 |
| Cit_clean/RD | 283,254 | 0.0079 | 0.4680 |
| Cit_dirty/RD | 283,254 | 0.0023 | 0.3460 |
| Cit_emtech/RD | 283,254 | 0.0263 | 0.9760 |
| Cit_other/RD | 283,254 | 0.2000 | 8.4510 |
| Firm traits | | | |
| RDG | 283,254 | 0.0377 | 0.1900 |
| invBE | 283,254 | -0.0078 | 0.9450 |
| taxRDBE | 283,254 | 0.1360 | 1.2250 |
| CEME | 283,254 | -0.0167 | 0.9440 |
| Earning˙abnormal | 283,254 | -0.0029 | 0.9390 |
| Adverts | 283,254 | 0.2570 | 2.6650 |

Notes. The Table presents summary statistics for Innovation productivity variables (RDBE, Pat/Book, Pat*/Book, Cit/Book and Cit*/Book), Innovation efficiency variables (Pat/RDC, Pat*/RDC, Cit/RD and Cit*/RD) and variables controlling for firm traits (Hirshleifer et al., 2013) during the period 1995-2012. The Variables are defined in Table 1.

**Table 3: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Intercept | 0.1930*** | 0.1950*** | 0.1950*** | 0.1920*** | 0.1950*** | 0.1950*** |
| | (0.0393) | (0.0393) | (0.0393) | (0.0392) | (0.0391) | (0.0391) |
| RDBE | 1.1330*** | 1.0820*** | 1.0730*** | 1.2690*** | 1.2570*** | 1.2580*** |
| | (0.0785) | (0.0781) | (0.0778) | (0.0822) | (0.0814) | (0.0814) |
| Pat/Book | 0.7190*** | | 0.2080 | | | |
| | (0.1230) | | (0.1080) | | | |
| Cit/Book | | 0.1740*** | 0.1460*** | | | |
| | | (0.0264) | (0.0276) | | | |
| Pat/RDC | | | | 0.0041* | | 0.0006 |
| | | | | (0.0017) | | (0.0007) |
| Cit/RD | | | | | 0.0147*** | 0.0146*** |
| | | | | | (0.0028) | (0.0028) |
| Time FE | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO | NO | NO |
| Observations | 79285 | 79285 | 79285 | 79284 | 79285 | 79284 |
| Adjusted $R^2$ | 0.2130 | 0.2150 | 0.2150 | 0.2090 | 0.2120 | 0.2120 |

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{i=1996}^{2012} \kappa_i year_i + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 4-6) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{i=2}^{17} \kappa_i year_i + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1-3 test whether the knowledge creation process acts as a continuum from R&D to patents to citations. And Models 4-6 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table 4: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.1950*** | 0.1950*** | 0.1950*** | 0.1950*** |
| | (0.0393) | (0.0393) | (0.0391) | (0.0391) |
| RDBE | 1.0720*** | 1.0720*** | 1.2580*** | 1.2560*** |
| | (0.0778) | (0.0779) | (0.0814) | (0.0813) |
| Pat_clean/Book | 1.8030** | | | |
| | (0.6150) | | | |
| Pat_dirty/Book | -0.9720 | | | |
| | (0.5520) | | | |
| Pat_other/Book | 0.1700 | | | |
| | (0.1090) | | | |
| Cit/Book | 0.1440*** | | | |
| | (0.0277) | | | |
| Cit_clean/Book | | 0.3220** | | |
| | | (0.1170) | | |
| Cit_dirty/Book | | -0.0876 | | |
| | | (0.1050) | | |
| Cit_other/Book | | 0.1390*** | | |
| | | (0.0291) | | |
| Pat/Book | | 0.2160* | | |
| | | (0.1080) | | |
| Pat_clean/RDC | | | 0.0588 | |
| | | | (0.0375) | |
| Pat_dirty/RDC | | | -0.0355** | |
| | | | (0.0137) | |
| Pat_other/RDC | | | 0.0005 | |
| | | | (0.0007) | |
| Cit/RD | | | 0.0144*** | |
| | | | (0.0028) | |
| Cit_clean/RD | | | | 0.0505* |
| | | | | (0.0236) |
| Cit_dirty/RD | | | | -0.0055 |
| | | | | (0.0048) |
| Cit_other/RD | | | | 0.0136*** |
| | | | | (0.0027) |
| Pat/RDC | | | | 0.0006 |
| | | | | (0.0007) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79285 | 79285 | 79284 | 79284 |
| Adjusted $R^2$ | 0.2150 | 0.2150 | 0.2120 | 0.2120 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table 5: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.0409* | 0.0410* | 0.0362 | 0.0365 |
| | (0.0219) | (0.0219) | (0.0228) | (0.0227) |
| RDBE | 0.2890*** | 0.2880*** | 0.3090*** | 0.3090*** |
| | (0.0267) | (0.0264) | (0.0365) | (0.0365) |
| Pat_clean/Book | 2.3190** | | | |
| | (0.9300) | | | |
| Pat_dirty/Book | -0.3290 | | | |
| | (1.2590) | | | |
| Pat_other/Book | 0.0690 | | | |
| | (0.0783) | | | |
| Cit/Book | 0.0324* | | | |
| | (0.0164) | | | |
| Cit_clean/Book | | 0.2890** | | |
| | | (0.1040) | | |
| Cit_dirty/Book | | -0.1150 | | |
| | | (0.2520) | | |
| Cit_other/Book | | 0.0321* | | |
| | | (0.0165) | | |
| Pat/Book | | 0.0770 | | |
| | | (0.0739) | | |
| Pat_clean/RDC | | | 0.1210** | |
| | | | (0.0455) | |
| Pat_dirty/RDC | | | 0.0181 | |
| | | | (0.0327) | |
| Pat_other/RDC | | | 0.0019* | |
| | | | (0.0011) | |
| Cit/RD | | | 0.0039*** | |
| | | | (0.0011) | |
| Cit_clean/RD | | | | 0.0224** |
| | | | | (0.0092) |
| Cit_dirty/RD | | | | -0.0099 |
| | | | | (0.0086) |
| Cit_other/RD | | | | 0.0042*** |
| | | | | (0.0012) |
| Pat/RDC | | | | 0.0025* |
| | | | | (0.0014) |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79,285 | 79,285 | 79,284 | 79,284 |
| avg. R-squared | 0.1930 | 0.1920 | 0.1880 | 0.1880 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table 6: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables for firms which conduct both clean and dirty innovation**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 1.4960*** | 1.5050*** | 1.4810*** | 1.4830*** |
| | (0.0173) | (0.0130) | (0.0156) | (0.0150) |
| RDBE | 0.0123 | 0.0033 | 0.0256 | 0.0240 |
| | (0.0171) | (0.0114) | (0.0150) | (0.0142) |
| Pat_clean/Book | 0.6160* | | | |
| | (0.2770) | | | |
| Pat_dirty/Book | -0.1030 | | | |
| | (0.2140) | | | |
| Pat_other/Book | -0.0394 | | | |
| | (0.0529) | | | |
| Cit/Book | 0.0238* | | | |
| | (0.0100) | | | |
| Cit_clean/Book | | 0.1240*** | | |
| | | (0.0221) | | |
| Cit_dirty/Book | | -0.0007 | | |
| | | (0.0088) | | |
| Cit_other/Book | | 0.0152 | | |
| | | (0.0082) | | |
| Pat/Book | | -0.0100 | | |
| | | (0.0291) | | |
| Pat_clean/RDC | | | 0.0026 | |
| | | | (0.0060) | |
| Pat_dirty/RDC | | | -0.0068* | |
| | | | (0.0033) | |
| Pat_other/RDC | | | -0.0017 | |
| | | | (0.0025) | |
| Cit/RD | | | 0.0013 | |
| | | | (0.0019) | |
| Cit_clean/RD | | | | 0.0179 |
| | | | | (0.0136) |
| Cit_dirty/RD | | | | -0.0031** |
| | | | | (0.0010) |
| Cit_other/RD | | | | -0.0003 |
| | | | | (0.0009) |
| Pat/RDC | | | | -0.0005 |
| | | | | (0.0008) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 6593 | 6593 | 6593 | 6593 |
| Adjusted $R^2$ | 0.2150 | 0.2180 | 0.1970 | 0.2040 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. In the above regression models the sample is the firms producing both clean and dirty technologies. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table 7: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.1950*** | 0.1950*** | 0.1950*** | 0.1960*** |
| | (0.0392) | (0.0393) | (0.0390) | (0.0391) |
| RDBE | 1.0710*** | 1.0690*** | 1.2520*** | 1.2420*** |
| | (0.0778) | (0.0777) | (0.0810) | (0.0807) |
| Pat_clean/Book | 1.7880** | | | |
| | (0.6050) | | | |
| Pat_dirty/Book | -0.9420 | | | |
| | (0.5670) | | | |
| Pat_emtech/Book | 0.6380 | | | |
| | (0.3550) | | | |
| Pat_other/Book | 0.0829 | | | |
| | (0.1070) | | | |
| Cit/Book | 0.1410*** | | | |
| | (0.0275) | | | |
| Cit_clean/Book | | 0.3160** | | |
| | | (0.1130) | | |
| Cit_dirty/Book | | -0.0820 | | |
| | | (0.1070) | | |
| Cit_emtech/Book | | 0.2490** | | |
| | | (0.0819) | | |
| Cit_other/Book | | 0.1150*** | | |
| | | (0.0332) | | |
| Pat/Book | | 0.2070 | | |
| | | (0.1080) | | |
| Pat_clean/RDC | | | 0.0459 | |
| | | | (0.0399) | |
| Pat_dirty/RDC | | | -0.0336* | |
| | | | (0.0131) | |
| Pat_emtech/RDC | | | 0.1950*** | |
| | | | (0.0431) | |
| Pat_other/RDC | | | 0.00003 | |
| | | | (0.00046) | |
| Cit/RD | | | 0.0123*** | |
| | | | (0.0027) | |
| Cit_clean/RD | | | | 0.0470* |
| | | | | (0.0227) |
| Cit_dirty/RD | | | | -0.0051 |
| | | | | (0.0048) |
| Cit_emtech/RD | | | | 0.0756*** |
| | | | | (0.0158) |
| Cit_other/RD | | | | 0.0082*** |
| | | | | (0.0024) |
| Pat/RDC | | | | 0.0005 |
| | | | | (0.0007) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79285 | 79285 | 79284 | 79284 |
| Adjusted $R^2$ | 0.2160 | 0.2150 | 0.2130 | 0.2130 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et. al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table 8: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, controlling for firm traits and emerging technology variants of Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Intercept | 0.2440*** | 0.2440*** | 0.2440*** | 0.2450*** | 0.2440*** | 0.2440*** | 0.2450*** | 0.2450*** |
| | (0.0331) | (0.0331) | (0.0329) | (0.0330) | (0.0331) | (0.0331) | (0.0329) | (0.0330) |
| RDBE | 0.3250*** | 0.3250*** | 0.2290*** | 0.2290*** | 0.3240*** | 0.3250*** | 0.2280*** | 0.2280*** |
| | (0.0361) | (0.0361) | (0.0274) | (0.0274) | (0.0359) | (0.0361) | (0.0274) | (0.0274) |
| Pat_clean/Book | 0.6490 | | | | 0.6480 | | | |
| | (0.4870) | | | | (0.4850) | | | |
| Pat_dirty/Book | 0.4910 | | | | 0.4940 | | | |
| | (0.9290) | | | | (0.9300) | | | |
| Pat_emtech/Book | | | | | 0.3050 | | | |
| | | | | | (0.2720) | | | |
| Pat_other/Book | 0.2210* | | | | 0.2070* | | | |
| | (0.0941) | | | | (0.0945) | | | |
| Cit/Book | 0.0620*** | | | | 0.0619*** | | | |
| | (0.0163) | | | | (0.0162) | | | |
| Cit_clean/Book | | 0.1440 | | | | 0.1440 | | |
| | | (0.0877) | | | | (0.0877) | | |
| Cit_dirty/Book | | -0.0151 | | | | -0.0151 | | |
| | | (0.0423) | | | | (0.0421) | | |
| Cit_emtech/Book | | | | | | 0.0585 | | |
| | | | | | | (0.0421) | | |
| Cit_other/Book | | 0.0595*** | | | | 0.0597** | | |
| | | (0.0167) | | | | (0.0195) | | |
| Pat/Book | | 0.2360* | | | | 0.2360* | | |
| | | (0.0936) | | | | (0.0939) | | |
| Pat_clean/RDC | | | 0.0422 | | | | 0.0316 | |
| | | | (0.0281) | | | | (0.0311) | |
| Pat_dirty/RDC | | | -0.0218 | | | | -0.0203 | |
| | | | (0.0146) | | | | (0.0140) | |
| Pat_emtech/RDC | | | | | | | 0.1440*** | |
| | | | | | | | (0.0313) | |
| Pat_other/RDC | | | 0.0010 | | | | 0.0002 | |
| | | | (0.0009) | | | | (0.0005) | |
| Cit/RD | | | 0.0066*** | | | | 0.0053*** | |
| | | | (0.0017) | | | | (0.0016) | |
| Cit_clean/RD | | | | 0.0446* | | | | 0.0429* |
| | | | | (0.0209) | | | | (0.0208) |
| Cit_dirty/RD | | | | -0.0016* | | | | -0.0015** |
| | | | | (0.0006) | | | | (0.0006) |
| Cit_emtech/RD | | | | | | | | 0.0394*** |
| | | | | | | | | (0.0097) |
| Cit_other/RD | | | | 0.0061*** | | | | 0.0033* |
| | | | | (0.0017) | | | | (0.0014) |
| Pat/RDC | | | | 0.0011 | | | | 0.0011 |
| | | | | (0.0009) | | | | (0.0009) |
| Time FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Firm-level controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 87800 | 87800 | 87799 | 87799 | 87800 | 87800 | 87799 | 87799 |
| Adjusted $R^2$ | 0.2480 | 0.2480 | 0.2440 | 0.2450 | 0.2480 | 0.2480 | 0.2450 | 0.2450 |

Notes. The Table presents the regression results of various specifications (columns 1, 2, 5 and 6) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3, 4, 7 and 8), $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3\gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005 and Hirshleifer et al., 2013 with the inclusion of firm-level control variables, year and industry fixed-effects. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process,from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table 9: Heckman sample selection $2^{nd}$ stage Model: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.2930*** | 0.2930*** | 0.4760*** | 0.4750*** |
| | (0.0641) | (0.0640) | (0.0575) | (0.0574) |
| RDBE | 0.5010*** | 0.4770*** | 0.7760*** | 0.7660*** |
| | (0.0443) | (0.0444) | (0.0408) | (0.0409) |
| PAT2c_book | 0.3290 | | 1.6270*** | |
| | (0.3790) | | (0.5020) | |
| PAT2d_book | 0.1860 | | 0.9730 | |
| | (1.2320) | | (1.1070) | |
| PAT2o_book | -0.1880*** | | -0.0957** | |
| | (0.0424) | | (0.0394) | |
| CITE2_book | 0.0409*** | | 0.0354*** | |
| | (0.0064) | | (0.0057) | |
| CITE2c_book | | 0.4490*** | | 0.4010*** |
| | | (0.0845) | | (0.0801) |
| CITE2d_book | | -0.1320 | | -0.0295 |
| | | (0.2450) | | (0.2210) |
| CITE2o_book | | 0.0340*** | | 0.0270*** |
| | | (0.0064) | | (0.0059) |
| PAT2_book | | -0.1830*** | | -0.0410 |
| | | (0.0307) | | (0.0377) |
| Inverse Mills Ratio | -.0941*** | -.0944*** | -.0526*** | -.0522*** |
| | (.0072 ) | (.0072) | (.0068) | (.0068) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | YES | YES |
| Observations | 78,577 | 78,577 | 78,577 | 78,577 |
| Censored observations | 68,103 | 68,103 | 68,103 | 68,103 |
| Uncensored observations | 10,474 | 10,474 | 10,474 | 10,474 |
| Wald Chiˆ2 | 3048.39 | 3076.36 | 6391.84 | 6406.00 |
| Prob > Chiˆ2 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Rho | -0.23333 | -0.23438 | -0.14714 | -0.14615 |
| Sigma | .40330519 | .40294769 | .35750088 | .35732054 |

Notes. The Table presents the regression results of various specifications of the $2^{nd}$ stage Heckman Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ The likelihood of a firm to conduct clean innovation is modeled in the $1^{st}$ stage of Heckman sample selection Model $Clean\_firm = \alpha + \gamma_1 Emtech\_firm_i + \gamma_2 RDG_{it} + \gamma_3 invBE_{it} + \gamma_4 taxRDBE_{it} + \gamma_5 CEME_{it} + \gamma_6 Earning_{abnormal_{it}} + \gamma_7 Adverts_{it} + \gamma_8 Total\_assets + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ where Clean_firm and Emtech_firm are indicator variables that take the value 1 if a firm has a USPTO published patent and 0 otherwise. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable in the $2^{nd}$ stage Model is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^{***}$ $p < 0.01$, $^{**}$ $p < 0.05$, $^{*}$ $p < 0.1$

## 2.8 Internet Appendices A-K

### 2.8.1 Internet Appendix A

**Table A1: Clean Patent classification codes**

| Patent code | Definition |
|---|---|
| Y02E | REDUCTION OF GREENHOUSE GAS [GHG] EMISSIONS, RELATED TO ENERGY GENERATION, TRANSMISSION OR DISTRIBUTION |
| Y02E10/00 | Energy generation through renewable energy sources |
| Y02E20/00 | Combustion technologies with mitigation potential |
| Y02E30/00 | Energy generation of nuclear origin |
| Y02E40/00 | Technologies for an efficient electrical power generation, transmission or distribution |
| Y02E50/00 | Technologies for the production of fuel of non-fossil origin |
| Y02E60/00 | Enabling technologies or technologies with a potential or indirect contribution to GHG emissions mitigation |
| Y02E70/00 | Other energy conversion or management systems reducing GHG emissions |
| Y02C | CAPTURE, STORAGE, SEQUESTRATION OR DISPOSAL OF GREENHOUSE GASES [GHG] |
| Y02C10/00 | CO2 capture or storage |
| Y02C20/00 | Capture or disposal of greenhouse gases [GHG] other than CO2 |
| Y02T | CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO TRANSPORTATION |
| Y02T10/00 | Road transport of goods or passengers |
| Y02T30/00 | Transportation of goods or passengers via railways |
| Y02T50/00 | Aeronautics or air transport |
| Y02T70/00 | Maritime or waterways transport |
| Y02T90/00 | Enabling technologies or technologies with a potential or indirect contribution to GHG emissions mitigation |
| Y02B | CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO BUILDINGS, e.g. HOUSING, HOUSE APPLIANCES OR RELATED END-USER APPLICATIONS |
| Y02B10/00 | Integration of renewable energy sources in buildings |
| Y02B20/00 | Energy efficient lighting technologies |
| Y02B30/00 | Energy efficient heating, ventilation or air conditioning [HVAC] |
| Y02B40/00 | Technologies aiming at improving the efficiency of home appliances |
| Y02B50/00 | Energy efficient technologies in elevators, escalators and moving walkways |
| Y02B70/00 | Technologies for an efficient end-user side electric power management and consumption |
| Y02B80/00 | Architectural or constructional elements improving the thermal performance of buildings |
| Y02B90/00 | Enabling technologies or technologies with a potential or indirect contribution to GHG emissions mitigation |

**Table A2: Dirty Patent classification codes**

| Patent code | Definition |
|---|---|
| C10J | PRODUCTION OF PRODUCER GAS, WATER-GAS, SYNTHESIS GAS FROM SOLID CARBONACEOUS MATERIAL, OR MIXTURES CONTAINING THESE GASES; CARBURETTING AIR OR OTHER GASES |
| F01K | STEAM ENGINE PLANTS; STEAM ACCUMULATORS; ENGINE PLANTS NOT OTHERWISE PROVIDED FOR; ENGINES USING SPECIAL WORKING FLUIDS OR CYCLES |
| F02C | GAS-TURBINE PLANTS; AIR INTAKES FOR JET-PROPULSION PLANTS; CONTROLLING FUEL SUPPLY IN AIR-BREATHING JET-PROPULSION PLANTS |
| F02G | HOT GAS OR COMBUSTION-PRODUCT POSITIVE-DISPLACEMENT ENGINE PLANTS; USE OF WASTE HEAT OF COMBUSTION ENGINES; NOT OTHERWISE PROVIDED FOR |
| F22 | STEAM GENERATION |
| F23 | COMBUSTION APPARATUS; COMBUSTION PROCESSES |
| F27 | FURNACES; KILNS; OVENS; RETORTS |
| F02B | INTERNAL-COMBUSTION PISTON ENGINES; COMBUSTION ENGINES IN GENERAL |

**Table A3: Emerging Technologies Patent classification codes**

| Patent code | Definition |
|---|---|
| Nanotechnology | |
| B82 | NANOTECHNOLOGY |
| GMO | |
| C12N/15 | MUTATION OR GENETIC ENGINEERING; DNA OR RNA CONCERNING GENETIC ENGINEERING, VECTORS, E.G., PLASMIDS, OR THEIR ISOLATION, PREPARATION OR PURIFICATION |
| 3D | |
| H04N/13 | STEREOSCOPIC VIDEO SYSTEMS; MULTI-VIEW VIDEO SYSTEMS |
| Wireless | |
| H04W | WIRELESS COMMUNICATION NETWORKS |
| Robots | |
| B25J | MANIPULATORS; CHAMBERS PROVIDED WITH MANIPULATION DEVICES |
| IT | |
| G06 (excl G06Q) | COMPUTING; CALCULATING; COUNTING |
| G10L | SPEECH ANALYSIS OR SYNTHESIS; SPEECH RECOGNITION; SPEECH OR VOICE PROCESSING; SPEECH OR AUDIO CODING OR DECODING |
| Biotechnology | |
| C07G | COMPOUNDS OF UNKNOWN CONSTITUTION |
| C07K | PEPTIDES |
| C12M | APPARATUS FOR ENZYMOLOGY OR MICROBIOLOGY |
| C12N | MICROORGANISMS OR ENZYMES; COMPOSITIONS THEREOF |
| C12P | FERMENTATION OR ENZYME-USING PROCESSES TO SYNTHESISE A DESIRED CHEMICAL COMPOUND OR COMPOSITION OR TO SEPARATE OPTICAL ISOMERS FROM A RACEMIC MIXTURE |
| C12Q | MEASURING OR TESTING PROCESSES INVOLVING ENZYMES, NUCLEIC ACIDS OR MICROORGANISMS; COMPOSITIONS OR TEST PAPERS THEREFOR; PROCESSES OF PREPARING SUCH COMPOSITIONS; CONDITION-RESPONSIVE CONTROL IN MICROBIOLOGICAL OR ENZYMOLOGICAL PROCESSES |
| C12R | MICROORGANISMS |

### 2.8.2 Internet Appendix B

**Table B1: Semi-elasticities for determining the impact of aggregated Innovation productivity and efficiency variables on Tobin's Q for the Models reported in Table 3**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RDBE | .8350 | .7970 | .7900 | .9362 | .9265 | .9267 |
| | (.0490) | (.0494) | (.0494) | (.0500) | (.0493) | (.0493) |
| Pat/Book | .5302 | | .1533 | | | |
| | (.0884) | | (.0790) | | | |
| Cit/Book | | .1281 | .1073 | | | |
| | | (.0188) | (.0199) | | | |
| Pat/RDC | | | | .0030 | | .0005 |
| | | | | (.0012) | | (.0005) |
| Cit/RD | | | | | .0109 | .0108 |
| | | | | | (.0020) | (.0020) |
| Observations | 79285 | 79285 | 79285 | 79284 | 79285 | 79284 |

Notes. The Table presents the semi-elasticities with respect to Innovation productivity and efficiency variables for various specifications (columns 1-3) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j)$ $+\varepsilon_{it}$ and the Model (columns 4-6) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} +$ $\sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors estimated using delta method in parentheses. All the variables are defined in Table 1.

**Table B2: Semi-elasticities for determining the impact of disaggregated Innovation productivity and efficiency variables on Tobin's Q for the Models reported in Table 4**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| RDBE | .7888 | .7889 | .9267 | .9253 |
| | (.0494) | ( .0494) | (.0493) | (.0493) |
| Pat_clean/Book | 1.3270 | | | |
| | (.4504) | | | |
| Pat_dirty/Book | -.7156 | | | |
| | (.4055) | | | |
| Pat_other/Book | .1251 | | | |
| | (.0800) | | | |
| Cit/Book | .1061 | | | |
| | (.0200) | | | |
| Cit_clean/Book | | .2370 | | |
| | | (.0854) | | |
| Cit_dirty/Book | | -.0645 | | |
| | | (.0770) | | |
| Cit_other/Book | | .1022 | | |
| | | (.0211) | | |
| Pat/Book | | .1588 | | |
| | | ( .0796) | | |
| Pat_clean/RDC | | | .0434 | |
| | | | (.0276) | |
| Pat_dirty/RDC | | | -.0261 | |
| | | | (.0101) | |
| Pat_other/RDC | | | .0004 | |
| | | | (.0005) | |
| Cit/RD | | | .0106 | |
| | | | (.0020 ) | |
| Cit_clean/RD | | | | .0372 |
| | | | | (.0173) |
| Cit_dirty/RD | | | | -.0041 |
| | | | | (.0035) |
| Cit_other/RD | | | | .0100 |
| | | | | (.0019) |
| Pat/RDC | | | | .0004 |
| | | | | (.0005) |
| Observations | 79285 | 79285 | 79284 | 79284 |

Notes. The Table presents the semi-elasticities with respect to Innovation productivity and efficiency variables for various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors estimated using delta method in parentheses. All the variables are defined in Table 1.

**Table B3: Semi-elasticities for determining the impact of disaggregated Innovation productivity and efficiency variables on Tobin's Q for the Models reported in Table 6**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| RDBE | .0381 | .0103 | .0773 | .0731 |
| | (.0525) | (.0354) | (.0443) | (.0423) |
| Pat_clean/Book | 1.9078 | | | |
| | (.8606) | | | |
| Pat_dirty/Book | -.3194 | | | |
| | (.6613) | | | |
| Pat_other/Book | -.1221 | | | |
| | (.1640) | | | |
| Cit/Book | .0738 | | | |
| | (.0312) | | | |
| Cit_clean/Book | | .3870 | | |
| | | (.0720) | | |
| Cit_dirty/Book | | -.0022 | | |
| | | (.0276) | | |
| Cit_other/Book | | .0476 | | |
| | | (.0259) | | |
| Pat/Book | | -.0311 | | |
| | | (.0910) | | |
| Pat_clean/RDC | | | .0078 | |
| | | | (.0180) | |
| Pat_dirty/RDC | | | -.0205 | |
| | | | (.0096) | |
| Pat_other/RDC | | | .0051 | |
| | | | (.0075) | |
| Cit/RD | | | .0038 | |
| | | | (.0056) | |
| Cit_clean/RD | | | | .0544 |
| | | | | (.0416) |
| Cit_dirty/RD | | | | -.0095 |
| | | | | (.0031) |
| Cit_other/RD | | | | -.0010 |
| | | | | (.0027) |
| Pat/RDC | | | | -.0016 |
| | | | | (.0025) |
| Observations | 6593 | 6593 | 6593 | 6593 |

Notes. The Table presents the semi-elasticities with respect to Innovation productivity and efficiency variables for various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors estimated using delta method in parentheses. All the variables are defined in Table 1.

**Table B4: Semi-elasticities for determining the impact of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables on Tobin's Q for the Models reported in Table 7**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| RDBE | .7891 | .7885 | .9254 | .9192 |
| | (.0494) | (.0494) | (.0493) | (.0492) |
| Pat_clean/Book | 1.3177 | | | |
| | (.4437) | | | |
| Pat_dirty/Book | -.6945 | | | |
| | (.4174 ) | | | |
| Pat_emtech/Book | .4700 | | | |
| | (.2611) | | | |
| Pat_other/Book | .0611 | | | |
| | (.0791) | | | |
| Cit/Book | .1043 | | | |
| | (.0199) | | | |
| Cit_clean/Book | | .2332 | | |
| | | ( .0831) | | |
| Cit_dirty/Book | | -.0604 | | |
| | | ( .0792) | | |
| Cit_emtech/Book | | .1837 | | |
| | | (.0601) | | |
| Cit_other/Book | | .0852 | | |
| | | (.0243) | | |
| Pat/Book | | .1528 | | |
| | | (.0793) | | |
| Pat_clean/RDC | | | .0339 | |
| | | | (.0295) | |
| Pat_dirty/RDC | | | -.0248 | |
| | | | (.0096) | |
| Pat_emtech/RDC | | | .1440 | |
| | | | (.0316) | |
| Pat_other/RDC | | | .00002 | |
| | | | (.00034) | |
| Cit/RD | | | .0091 | |
| | | | ( .0019) | |
| Cit_clean/RD | | | | .0348 |
| | | | | (.0167) |
| Cit_dirty/RD | | | | -.0038 |
| | | | | (.0036) |
| Cit_emtech/RD | | | | .0560 |
| | | | | (.0116) |
| Cit_other/RD | | | | .0061 |
| | | | | (.0018) |
| Pat/RDC | | | | .0004 |
| | | | | (.0005) |
| Observations | 79,285 | 79,285 | 79,284 | 79,284 |

Notes. The Table presents the semi-elasticities with respect to Innovation productivity and efficiency variables for various specifications (columns 1-2) of the Model, $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4), $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors estimated using delta method in parentheses. All the variables are defined in Table 1.

**Table B5: Semi-elasticities for determining the impact of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables on Tobin's Q for the Models reported in Table 8**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| RDBE | .2976 | .2978 | .2150 | .2148 | .2966 | .2978 | .2137 | .2134 |
| | (.0311) | (.0311) | (.0245) | (.0245) | (.0310) | (.0311) | (.0244) | (.0245) |
| Pat_clean/Book | .5948 | | | | .5940 | | | |
| | (.4461) | | | | (.4439) | | | |
| Pat_dirty/Book | .4501 | | | | .4525 | | | |
| | (.8513) | | | | (.8524) | | | |
| Pat_emtech/Book | | | | | .2797 | | | |
| | | | | | (.2480) | | | |
| Pat_other/Book | .2027 | | | | .1895 | | | |
| | (.0857) | | | | (.0862) | | | |
| Cit/Book | .0568 | | | | .0567 | | | |
| | (.0148) | | | | (.0147) | | | |
| Cit_clean/Book | | .1316 | | | | .1317 | | |
| | | (.0803) | | | | (.0803) | | |
| Cit_dirty/Book | | -.0139 | | | | -.0139 | | |
| | | (.0387) | | | | (.0386) | | |
| Cit_emtech/Book | | | | | | .0536 | | |
| | | | | | | (.0385) | | |
| Cit_other/Book | | .0545 | | | | .0547 | | |
| | | (.0152) | | | | (.0178) | | |
| Pat/Book | | .2161 | | | | .2161 | | |
| | | (.0851) | | | | (.0854) | | |
| Pat_clean/RDC | | | .0396 | | | | .0296 | |
| | | | (.0263) | | | | (.0291) | |
| Pat_dirty/RDC | | | -.0204 | | | | -.0190 | |
| | | | (.0137) | | | | (.0131) | |
| Pat_emtech/RDC | | | | | | | .1351 | |
| | | | | | | | (.0289) | |
| Pat_other/RDC | | | .0009 | | | | .0002 | |
| | | | (.0008) | | | | (.0005) | |
| Cit/RD | | | .0062 | | | | .0050 | |
| | | | (.0016) | | | | (.0015) | |
| Cit_clean/RD | | | | .0418 | | | | .0401 |
| | | | | (.0195) | | | | (.0194) |
| Cit_dirty/RD | | | | -.0015 | | | | -.0014 |
| | | | | (.0006) | | | | (.0006) |
| Cit_emtech/RD | | | | | | | | .0369 |
| | | | | | | | | (.0090) |
| Cit_other/RD | | | | .0058 | | | | .0031 |
| | | | | (.0015) | | | | (.0013) |
| Pat/RDC | | | | .0010 | | | | .0010 |
| | | | | (.0008) | | | | (.0008) |
| Observations | 87,800 | 87,800 | 87,799 | 87,799 | 87,800 | 87,800 | 87,799 | 87,799 |

Notes. The Table presents the semi-elasticities with respect to Innovation productivity and efficiency variables for various specifications (columns 1, 2, 5 and 6) of the Model, $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3, 4, 7 and 8), $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors estimated using delta method in parentheses. All the variables are defined in Table 1.

**Table B6: The Table reports the nonlinear hypothesis for the coefficients from the Models reported in Tables 4, 6, 7, and 8**

| t Test for Non-linear least squares estimation | $Chi^2$ | $Prob > Chi^2$ |
|---|---|---|
| Pat_clean/Book - Pat_dirty/Book = 0 | 8.58 | 0.0034 |
| Cit_clean/Book - Cit_dirty/Book = 0 | 6.28 | 0.0122 |
| Pat_clean/RDC - Pat_dirty/RDC = 0 | 6.36 | 0.0116 |
| Cit_clean/RD - Cit_dirty/RD = 0 | 5.81 | 0.0160 |

**(a) Panel A: Test for Non-linear hypotheses after estimation for the Models reported in Table 4**

| t Test for Non-linear least squares estimation | $Chi^2$ | $Prob > Chi^2$ |
|---|---|---|
| Pat_clean/Book - Pat_dirty/Book = 0 | 3.75 | 0.0528 |
| Cit_clean/Book - Cit_dirty/Book = 0 | 26.19 | 0.0000 |
| Pat_clean/RDC - Pat_dirty/RDC = 0 | 2.50 | 0.1141 |
| Cit_clean/RD - Cit_dirty/RD = 0 | 2.38 | 0.1233 |

**(b) Panel B: Test for Non-linear hypotheses after estimation for the Models reported in Table 6**

| t Test for Non-linear least squares estimation | $Chi^2$ | $Prob > Chi^2$ |
|---|---|---|
| Pat_clean/Book - Pat_dirty/Book = 0 | 8.38 | 0.0038 |
| Cit_clean/Book - Cit_dirty/Book = 0 | 5.99 | 0.0144 |
| Pat_clean/RDC - Pat_dirty/RDC = 0 | 3.44 | 0.0638 |
| Cit_clean/RD - Cit_dirty/RD = 0 | 5.41 | 0.0201 |

**(c) Panel C: Test for Non-linear hypotheses after estimation for the Models reported in Table 7**

| t Test for Non-linear least squares estimation | $Chi^2$ | $Prob > Chi^2$ |
|---|---|---|
| Pat_clean/Book - Pat_dirty/Book = 0 | 0.02 | 0.8822 |
| Cit_clean/Book - Cit_dirty/Book = 0 | 2.62 | 0.1057 |
| Pat_clean/RDC - Pat_dirty/RDC = 0 | 5.74 | 0.0166 |
| Cit_clean/RD - Cit_dirty/RD = 0 | 4.90 | 0.0268 |

**(d) Panel D:Test for Non-linear hypotheses after estimation for the Models (1-4) reported in Table 8**

| t Test for Non-linear least squares estimation | $Chi^2$ | $Prob > Chi^2$ |
|---|---|---|
| Pat_clean/Book - Pat_dirty/Book = 0 | 0.02 | 0.8848 |
| Cit_clean/Book - Cit_dirty/Book = 0 | 2.62 | 0.1054 |
| Pat_clean/RDC - Pat_dirty/RDC = 0 | 2.78 | 0.0957 |
| Cit_clean/RD - Cit_dirty/RD = 0 | 4.57 | 0.0326 |

**(e) Panel E: Test for Non-linear hypotheses after estimation for the Models (5-8) reported in Table 8**

### 2.8.3 Internet Appendix C

**Table C1: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Intercept | 0.2420*** | 0.2390*** | 0.2410*** | 0.2440*** | 0.2430*** | 0.2440*** |
| | (0.0361) | (0.0359) | (0.0360) | (0.0369) | (0.0370) | (0.0369) |
| RDBE | 0.6480*** | 0.8090*** | 0.7160*** | 0.261*** | 0.2600*** | 0.2610*** |
| | (0.0432) | (0.0545) | (0.0502) | (0.0199) | (0.0199) | (0.0199) |
| Pat/Book | 1.9310*** | | 1.5550*** | | | |
| | (0.1860) | | (0.2410) | | | |
| Cit/Book | | 0.1060*** | 0.0282* | | | |
| | | (0.0121) | (0.0139) | | | |
| Pat/RDC | | | | 0.0262*** | | 0.0255*** |
| | | | | (0.0059) | | (0.0058) |
| Cit/RD | | | | | 0.0016 | 0.0007 |
| | | | | | (0.0009) | (0.0008) |
| Time FE | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO | NO | NO |
| Observations | 87800 | 87800 | 87800 | 87800 | 87800 | 87800 |
| Adjusted $R^2$ | 0.2040 | 0.2000 | 0.2040 | 0.1900 | 0.1900 | 0.1900 |

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 4-6) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1-3 test whether the knowledge creation process acts as a continuum from R&D to patents to citations. And Models 4-6 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table C2: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.2410*** | 0.2410*** | 0.2440*** | 0.2440*** |
| | (0.0360) | (0.0360) | (0.0369) | (0.0369) |
| RDBE | 0.7150*** | 0.7300*** | 0.2610*** | 0.2610*** |
| | (0.0503) | (0.0514) | (0.0199) | (0.0199) |
| Pat_clean/Book | 3.3150* | | | |
| | (1.3620) | | | |
| Pat_dirty/Book | 0.8690 | | | |
| | (1.1040) | | | |
| Pat_other/Book | 1.5020*** | | | |
| | (0.2430) | | | |
| Cit/Book | 0.0301* | | | |
| | (0.0139) | | | |
| Cit_clean/Book | | 0.0011 | | |
| | | (0.0024) | | |
| Cit_dirty/Book | | -0.0624*** | | |
| | | (0.0102) | | |
| Cit_other/Book | | 0.0330* | | |
| | | (0.0148) | | |
| Pat/Book | | 1.5030*** | | |
| | | (0.2430) | | |
| Pat_clean/RDC | | | 0.0304 | |
| | | | (0.0239) | |
| Pat_dirty/RDC | | | -0.0336 | |
| | | | (0.0448) | |
| Pat_other/RDC | | | 0.0259*** | |
| | | | (0.0062) | |
| Cit/RD | | | 0.0008 | |
| | | | (0.0008) | |
| Cit_clean/RD | | | | 0.0004 |
| | | | | (0.0025) |
| Cit_dirty/RD | | | | -0.0011*** |
| | | | | (0.0001) |
| Cit_other/RD | | | | 0.0011 |
| | | | | (0.0010) |
| Pat/RDC | | | | 0.0253*** |
| | | | | (0.0057) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 87800 | 87800 | 87800 | 87800 |
| Adjusted $R^2$ | 0.2040 | 0.2040 | 0.1900 | 0.1900 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table C3: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.2410*** | 0.2410*** | 0.2440*** | 0.2440*** |
| | (0.0360) | (0.0360) | (0.0369) | (0.0369) |
| RDBE | 0.7140*** | 0.7360*** | 0.2600*** | 0.261*** |
| | (0.0503) | (0.0524) | (0.0199) | (0.0199) |
| Pat_clean/Book | 3.2350* | | | |
| | (1.3590) | | | |
| Pat_dirty/Book | 0.8810 | | | |
| | (1.1200) | | | |
| Pat_emtech/Book | 2.2050*** | | | |
| | (0.6400) | | | |
| Pat_other/Book | 1.3570*** | | | |
| | (0.2540) | | | |
| Cit/Book | 0.0354* | | | |
| | (0.0144) | | | |
| Cit_clean/Book | | 0.0012 | | |
| | | (0.0024) | | |
| Cit_dirty/Book | | -0.0623*** | | |
| | | (0.0101) | | |
| Cit_emtech/Book | | 0.0029 | | |
| | | (0.0304) | | |
| Cit_other/Book | | 0.0345* | | |
| | | (0.0152) | | |
| Pat/Book | | 1.4970*** | | |
| | | (0.2430) | | |
| Pat_clean/RDC | | | 0.0282 | |
| | | | (0.0218) | |
| Pat_dirty/RDC | | | -0.0265 | |
| | | | (0.0435) | |
| Pat_emtech/RDC | | | 0.1660** | |
| | | | (0.0630) | |
| Pat_other/RDC | | | 0.0196*** | |
| | | | (0.0058) | |
| Cit/RD | | | 0.0003 | |
| | | | (0.0008) | |
| Cit_clean/RD | | | | 0.0003 |
| | | | | (0.0024) |
| Cit_dirty/RD | | | | -0.0011*** |
| | | | | (0.0001) |
| Cit_emtech/RD | | | | 0.0016* |
| | | | | (0.0008) |
| Cit_other/RD | | | | 0.0008 |
| | | | | (0.0016) |
| Pat/RDC | | | | 0.0253*** |
| | | | | (0.0057) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 87800 | 87800 | 87800 | 87800 |
| Adjusted $R^2$ | 0.2040 | 0.2040 | 0.1910 | 0.1900 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model, $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4), $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ , that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table C4: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Intercept | 0.0381 | 0.0382 | 0.0381 | 0.0381 | 0.0381 | 0.0381 |
| | (0.0249) | (0.0249) | (0.0249) | (0.0249) | (0.0249) | (0.0249) |
| RDBE | 0.0867*** | 0.0868*** | 0.0860*** | 0.0879*** | 0.0878*** | 0.0879*** |
| | (0.0177) | (0.0188) | (0.0177) | (0.0189) | (0.0189) | (0.0189) |
| Pat/Book | 0.3430*** | | 0.4060*** | | | |
| | (0.1150) | | (0.1350) | | | |
| Cit/Book | | 0.0251*** | -0.0020 | | | |
| | | (0.0060) | (0.0103) | | | |
| Pat/RDC | | | | 0.0089*** | | 0.0095*** |
| | | | | (0.0022) | | (0.0023) |
| Cit/RD | | | | | 0.0004 | -9.45e-05 |
| | | | | | (0.0005) | (0.0005) |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO | NO | NO |
| Observations | 87,800 | 87,800 | 87,800 | 87,800 | 87,800 | 87,800 |
| avg. R-squared | 0.1810 | 0.1770 | 0.1820 | 0.1770 | 0.1760 | 0.1770 |

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 4-6) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to patents and citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table C5: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.0381 | 0.0381 | 0.0381 | 0.0381 |
| | (0.0249) | (0.0249) | (0.0249) | (0.0249) |
| RDBE | 0.0860*** | 0.0862*** | 0.0879*** | 0.0879*** |
| | (0.0177) | (0.0177) | (0.0188) | (0.0189) |
| Pat_clean/Book | 4.1090*** | | | |
| | (1.3970) | | | |
| Pat_dirty/Book | 0.7770 | | | |
| | (0.9080) | | | |
| Pat_other/Book | 0.4100*** | | | |
| | (0.1340) | | | |
| Cit/Book | -0.0029 | | | |
| | (0.0101) | | | |
| Cit_clean/Book | | 0.1970 | | |
| | | (0.1430) | | |
| Cit_dirty/Book | | -0.2280 | | |
| | | (0.2300) | | |
| Cit_other/Book | | -0.0013 | | |
| | | (0.0107) | | |
| Pat/Book | | 0.4210*** | | |
| | | (0.1380) | | |
| Pat_clean/RDC | | | 0.1520* | |
| | | | (0.0855) | |
| Pat_dirty/RDC | | | 0.0870 | |
| | | | (0.0954) | |
| Pat_other/RDC | | | 0.0102*** | |
| | | | (0.0029) | |
| Cit/RD | | | -1.04e-06 | |
| | | | (0.0006) | |
| Cit_clean/RD | | | | -0.0093 |
| | | | | (0.0069) |
| Cit_dirty/RD | | | | -0.0145 |
| | | | | (0.0161) |
| Cit_other/RD | | | | 0.0002 |
| | | | | (0.0006) |
| Pat/RDC | | | | 0.0101*** |
| | | | | (0.0027) |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 87,800 | 87,800 | 87,800 | 87,800 |
| avg. R-squared | 0.1830 | 0.1830 | 0.1780 | 0.1770 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table C6: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.0380 | 0.0380 | 0.0380 | 0.0381 |
| | (0.0249) | (0.0249) | (0.0249) | (0.0249) |
| RDBE | 0.0883*** | 0.0870*** | 0.0879*** | 0.0879*** |
| | (0.0171) | (0.0176) | (0.0188) | (0.0189) |
| Pat_clean/Book | 4.0970*** | | | |
| | (1.3810) | | | |
| Pat_dirty/Book | 0.6930 | | | |
| | (0.9070) | | | |
| Pat_emtech/Book | 0.7940** | | | |
| | (0.3010) | | | |
| Pat_other/Book | 0.3890** | | | |
| | (0.1380) | | | |
| Cit/Book | 0.0033 | | | |
| | (0.0106) | | | |
| Cit_clean/Book | | 0.1860 | | |
| | | (0.1420) | | |
| Cit_dirty/Book | | -0.2280 | | |
| | | (0.2300) | | |
| Cit_emtech/Book | | 0.0959** | | |
| | | (0.0403) | | |
| Cit_other/Book | | -0.0111 | | |
| | | (0.0107) | | |
| Pat/Book | | 0.4310*** | | |
| | | (0.1390) | | |
| Pat_clean/RDC | | | 0.1450* | |
| | | | (0.0823) | |
| Pat_dirty/RDC | | | 0.0922 | |
| | | | (0.0961) | |
| Pat_emtech/RDC | | | 0.0750** | |
| | | | (0.0276) | |
| Pat_other/RDC | | | 0.0097*** | |
| | | | (0.0031) | |
| Cit/RD | | | -0.00027 | |
| | | | (0.000612) | |
| Cit_clean/RD | | | | -0.0106 |
| | | | | (0.0073) |
| Cit_dirty/RD | | | | -0.0140 |
| | | | | (0.0151) |
| Cit_emtech/RD | | | | 0.0028** |
| | | | | (0.0012) |
| Cit_other/RD | | | | -0.0003 |
| | | | | (0.0007) |
| Pat/RDC | | | | 0.0103*** |
| | | | | (0.0028) |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 87,800 | 87,800 | 87,800 | 87,800 |
| avg. R-squared | 0.1850 | 0.1830 | 0.1780 | 0.1770 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model, $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table C7: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, controlling for firm traits and emerging technology variants of Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Intercept | 0.1610*** | 0.1610*** | 0.1610*** | 0.1610*** | 0.1610*** | 0.1610*** | 0.1610*** | 0.1610*** |
| | (0.0319) | (0.0319) | (0.0319) | (0.0319) | (0.0319) | (0.0319) | (0.0319) | (0.0319) |
| RDBE | 0.0550*** | 0.0553*** | 0.0610*** | 0.0610*** | 0.0569*** | 0.0561*** | 0.0610*** | 0.0610*** |
| | (0.0145) | (0.0146) | (0.0155) | (0.0155) | (0.0144) | (0.0146) | (0.0155) | (0.0155) |
| Pat_clean/Book | 3.4130*** | | | | 3.4330*** | | | |
| | (1.0530) | | | | (1.0430) | | | |
| Pat_dirty/Book | 0.7030 | | | | 0.6430 | | | |
| | (0.7740) | | | | (0.7680) | | | |
| Pat_emtech/Book | | | | | 0.8110*** | | | |
| | | | | | (0.1950) | | | |
| Pat_other/Book | 0.3880*** | | | | 0.3850*** | | | |
| | (0.1160) | | | | (0.1220) | | | |
| Cit/Book | -0.0040 | | | | -0.0005 | | | |
| | (0.0103) | | | | (0.0105) | | | |
| Cit_clean/Book | | 0.1910 | | | | 0.1810 | | |
| | | (0.1360) | | | | (0.1340) | | |
| Cit_dirty/Book | | 0.0905 | | | | 0.0917 | | |
| | | (0.2450) | | | | (0.2450) | | |
| Cit_emtech/Book | | | | | | 0.0719* | | |
| | | | | | | (0.0397) | | |
| Cit_other/Book | | -0.0033 | | | | -0.0111 | | |
| | | (0.0110) | | | | (0.0103) | | |
| Pat/Book | | 0.3970*** | | | | 0.4070*** | | |
| | | (0.1200) | | | | (0.1200) | | |
| Pat_clean/RDC | | | 0.1190* | | | | 0.1140* | |
| | | | (0.0657) | | | | (0.0634) | |
| Pat_dirty/RDC | | | 0.0924 | | | | 0.0962 | |
| | | | (0.0810) | | | | (0.0816) | |
| Pat_emtech/RDC | | | | | | | 0.0581** | |
| | | | | | | | (0.0220) | |
| Pat_other/RDC | | | 0.0092*** | | | | 0.0092*** | |
| | | | (0.0025) | | | | (0.0029) | |
| Cit/RD | | | -3.31e-05 | | | | -0.0003 | |
| | | | (0.0006) | | | | (0.0006) | |
| Cit_clean/RD | | | | -0.0110 | | | | -0.0117 |
| | | | | (0.0078) | | | | (0.0081) |
| Cit_dirty/RD | | | | 0.0062 | | | | 0.0074 |
| | | | | (0.0104) | | | | (0.0107) |
| Cit_emtech/RD | | | | | | | | -0.0005 |
| | | | | | | | | (0.0014) |
| Cit_other/RD | | | | 2.31e-05 | | | | -0.0003 |
| | | | | (0.0006) | | | | (0.0007) |
| Pat/RDC | | | | 0.0091*** | | | | 0.0092*** |
| | | | | (0.0023) | | | | (0.0024) |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Firm-level controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 87,800 | 87,800 | 87,800 | 87,800 | 87,800 | 87,800 | 87,800 | 87,800 |
| avg. R-squared | 0.3030 | 0.3030 | 0.2990 | 0.2980 | 0.3040 | 0.3030 | 0.2990 | 0.2980 |

Notes. The Table presents the regression results of various specifications (columns 1, 2, 5 and 6) of the Model, $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it}$ $+ \gamma_8 Earningabnormalit + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 2, 4, 7 and 8) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}}$ $+\gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

### 2.8.4 Internet Appendix D

**Table D1: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Intercept | 0.0409* | 0.0411* | 0.0411* | 0.0405* | 0.0372 | 0.0362 |
| | (0.0219) | (0.0219) | (0.0219) | (0.0218) | (0.0226) | (0.0228) |
| RDBE | 0.2850*** | 0.2800*** | 0.2840*** | 0.3090*** | 0.3080*** | 0.3090*** |
| | (0.0288) | (0.0284) | (0.0280) | (0.0364) | (0.0364) | (0.0364) |
| Pat/Book | 0.1740*** | | 0.0712 | | | |
| | (0.0469) | | (0.0676) | | | |
| Cit/Book | | 0.0390*** | 0.0341** | | | |
| | | (0.0113) | (0.0159) | | | |
| Pat/RDC | | | | 0.0044*** | | 0.0024 |
| | | | | (0.0015) | | (0.0014) |
| Cit/RD | | | | | 0.0037*** | 0.0037*** |
| | | | | | (0.0011) | (0.0011) |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO | NO | NO |
| Observations | 79,285 | 79,285 | 79,285 | 79,284 | 79,285 | 79,284 |
| avg. R-squared | 0.1870 | 0.1890 | 0.1900 | 0.1850 | 0.1870 | 0.1870 |

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 4-6) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to patents and citations and tests the efficiency in the knowledge creation process,from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table D2: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.0410* | 0.0410* | 0.0363 | 0.0373 |
| | (0.0219) | (0.0220) | (0.0226) | (0.0225) |
| RDBE | 0.2880*** | 0.2880*** | 0.3080*** | 0.3080*** |
| | (0.0266) | (0.0258) | (0.0364) | (0.0365) |
| Pat_clean/Book | 2.3010** | | | |
| | (0.9200) | | | |
| Pat_dirty/Book | -0.3480 | | | |
| | (1.2570) | | | |
| Pat_emtech/Book | 0.3090* | | | |
| | (0.1600) | | | |
| Pat_other/Book | 0.0422 | | | |
| | (0.0724) | | | |
| Cit/Book | 0.0357** | | | |
| | (0.0158) | | | |
| Cit_clean/Book | | 0.2860** | | |
| | | (0.1000) | | |
| Cit_dirty/Book | | -0.1170 | | |
| | | (0.2520) | | |
| Cit_emtech/Book | | 0.1640*** | | |
| | | (0.0486) | | |
| Cit_other/Book | | 0.0268 | | |
| | | (0.0161) | | |
| Pat/Book | | 0.0461 | | |
| | | (0.0707) | | |
| Pat_clean/RDC | | | 0.1070** | |
| | | | (0.0478) | |
| Pat_dirty/RDC | | | 0.0246 | |
| | | | (0.0314) | |
| Pat_emtech/RDC | | | 0.0787*** | |
| | | | (0.0181) | |
| Pat_other/RDC | | | -0.0005 | |
| | | | (0.0011) | |
| Cit/RD | | | 0.0036*** | |
| | | | (0.0010) | |
| Cit_clean/RD | | | | 0.0240** |
| | | | | (0.0084) |
| Cit_dirty/RD | | | | -0.0095 |
| | | | | (0.0085) |
| Cit_emtech/RD | | | | 0.0278** |
| | | | | (0.0096) |
| Cit_other/RD | | | | 0.0026** |
| | | | | (0.0009) |
| Pat/RDC | | | | 0.0025* |
| | | | | (0.0014) |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79,285 | 79,285 | 79,284 | 79,284 |
| avg. R-squared | 0.1940 | 0.1940 | 0.1890 | 0.1900 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table D3: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, controlling for firm traits and emerging technology variants of Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Intercept | 0.1610*** | 0.1610*** | 0.1590*** | 0.1590*** | 0.1610*** | 0.1610*** | 0.1600*** | 0.1590*** |
| | (0.0320) | (0.0320) | (0.0323) | (0.0323) | (0.0320) | (0.0320) | (0.0322) | (0.0323) |
| RDBE | 0.0569*** | 0.0516*** | 0.0609*** | 0.0609*** | 0.0593*** | 0.0537*** | 0.0608*** | 0.0609*** |
| | (0.0121) | (0.0120) | (0.0155) | (0.0155) | (0.0124) | (0.0128) | (0.0155) | (0.0155) |
| Pat_clean/Book | 1.5640* | | | | 1.5490** | | | |
| | (0.7390) | | | | (0.7260) | | | |
| Pat_dirty/Book | 0.2340 | | | | 0.0946 | | | |
| | (1.2170) | | | | (1.1950) | | | |
| Pat_emtech/Book | | | | | 0.2910** | | | |
| | | | | | (0.1160) | | | |
| Pat_other/Book | 0.1400*** | | | | 0.1690** | | | |
| | (0.0670) | | | | (0.0595) | | | |
| Cit/Book | 0.0160 | | | | 0.0146* | | | |
| | (0.0094) | | | | (0.0078) | | | |
| Cit_clean/Book | | 0.2090** | | | | 0.2050** | | |
| | | (0.0979) | | | | (0.0975) | | |
| Cit_dirty/Book | | 0.0775 | | | | 0.0688 | | |
| | | (0.1790) | | | | (0.1750) | | |
| Cit_emtech/Book | | | | | | 0.0301** | | |
| | | | | | | (0.0138) | | |
| Cit_other/Book | | 0.0153 | | | | 0.0158 | | |
| | | (0.0093) | | | | (0.0095) | | |
| Pat/Book | | 0.1420** | | | | 0.1430** | | |
| | | (0.0638) | | | | (0.0599) | | |
| Pat_clean/RDC | | | 0.0976** | | | | 0.0888** | |
| | | | (0.0393) | | | | (0.0397) | |
| Pat_dirty/RDC | | | 0.0169 | | | | 0.0208 | |
| | | | (0.0421) | | | | (0.0420) | |
| Pat_emtech/RDC | | | | | | | 0.0499*** | |
| | | | | | | | (0.0132) | |
| Pat_other/RDC | | | 0.0028*** | | | | 0.0008 | |
| | | | (0.0009) | | | | (0.0009) | |
| Cit/RD | | | 0.0015*** | | | | 0.0013*** | |
| | | | (0.0004) | | | | (0.0003) | |
| Cit_clean/RD | | | | 0.0172** | | | | 0.0177** |
| | | | | (0.0065) | | | | (0.0063) |
| Cit_dirty/RD | | | | 0.0036 | | | | 0.0039 |
| | | | | (0.0109) | | | | (0.0110) |
| Cit_emtech/RD | | | | | | | | 0.0110*** |
| | | | | | | | | (0.0037) |
| Cit_other/RD | | | | 0.0023** | | | | 0.0019*** |
| | | | | (0.0008) | | | | (0.0006) |
| Pat/RDC | | | | 0.0031** | | | | 0.0029** |
| | | | | (0.0011) | | | | (0.0012) |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Firm-level controls | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 87,800 | 87,800 | 87,799 | 87,799 | 87,800 | 87,800 | 87,799 | 87,799 |
| avg. R-squared | 0.3050 | 0.3050 | 0.2990 | 0.2990 | 0.3060 | 0.3060 | 0.3000 | 0.3000 |

Notes. The Table presents the regression results of various specifications (columns 1, 2, 5 and 6) of the Model, $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 2, 4, 7 and 8) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_6 invBE_{it} + \gamma_5 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earning_{abnormal_{it}} + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

### 2.8.5 Internet Appendix E: Operating performance

**Table E1: Subsequent year's EBITDA as a function of disaggregated Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 7.821*** | 7.771*** | 7.813*** | 7.805*** |
| | (0.227) | (0.240) | (0.225) | (0.226) |
| RDBE | -0.0529*** | -0.0531** | -0.0516*** | -0.0521*** |
| | (0.0152) | (0.0163) | (0.0151) | (0.0152) |
| Pat_clean/Book | 1458.2*** | | | |
| | (388.7) | | | |
| Pat_dirty/Book | -0.956** | | | |
| | (0.368) | | | |
| Pat_other/Book | -0.0481 | | | |
| | (0.0469) | | | |
| Cit/Book | -0.00682*** | | | |
| | (0.00173) | | | |
| Cit_clean/Book | | -0.177*** | | |
| | | (0.0511) | | |
| Cit_dirty/Book | | 653.7** | | |
| | | (223.4) | | |
| Cit_other/Book | | -0.0189** | | |
| | | (0.00589) | | |
| Pat/Book | | 0.961* | | |
| | | (0.379) | | |
| Pat_clean/RDC | | | 111.8*** | |
| | | | (30.36) | |
| Pat_dirty/RDC | | | -0.0257** | |
| | | | (0.00917) | |
| Pat_other/RDC | | | -0.000262 | |
| | | | (0.00260) | |
| Cit/RD | | | 0.00140 | |
| | | | (0.00231) | |
| Cit_clean/RD | | | | 29.02*** |
| | | | | (8.071) |
| Cit_dirty/RD | | | | -0.0120* |
| | | | | (0.00468) |
| Cit_other/RD | | | | 0.000246 |
| | | | | (0.00158) |
| Pat/RDC | | | | 0.000325 |
| | | | | (0.00282) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 63071 | 63071 | 63070 | 63070 |
| Adjusted $R^2$ | 0.203 | 0.197 | 0.197 | 0.196 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logEBITDA_{i,t+1} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logEBITDA_{i,t+1} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method. These Models assess the relationship between a firm's operating performance, measured by EBITDA, in year $t+1$ with Innovation productivity and efficiency variables in year $t$. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of EBITDA and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table E2: Subsequent year's EBITDA as a function of disaggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 7.604*** | 7.602*** | 7.600*** | 7.599*** |
| | (0.0577) | (0.0585) | (0.0586) | (0.0589) |
| RDBE | 0.651* | 0.636* | 0.591 | 0.591 |
| | (0.348) | (0.348) | (0.341) | (0.341) |
| Pat_clean/Book | 10.51** | | | |
| | (4.300) | | | |
| Pat_dirty/Book | 5.418 | | | |
| | (6.384) | | | |
| Pat_other/Book | 0.151 | | | |
| | (0.229) | | | |
| Cit/Book | -0.130*** | | | |
| | (0.0425) | | | |
| Cit_clean/Book | | 0.777 | | |
| | | (0.585) | | |
| Cit_dirty/Book | | 3.836*** | | |
| | | (1.235) | | |
| Cit_other/Book | | -0.146*** | | |
| | | (0.0442) | | |
| Pat/Book | | 0.186 | | |
| | | (0.205) | | |
| Pat_clean/RDC | | | 0.503** | |
| | | | (0.179) | |
| Pat_dirty/RDC | | | -0.00213 | |
| | | | (0.202) | |
| Pat_other/RDC | | | -0.00742 | |
| | | | (0.00468) | |
| Cit/RD | | | 0.00342 | |
| | | | (0.00375) | |
| Cit_clean/RD | | | | 0.128** |
| | | | | (0.0566) |
| Cit_dirty/RD | | | | -0.0844 |
| | | | | (0.0740) |
| Cit_other/RD | | | | 0.00276 |
| | | | | (0.00294) |
| Pat/RDC | | | | -0.00493 |
| | | | | (0.00357) |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 63,071 | 63,071 | 63,070 | 63,070 |
| avg. R-squared | 0.220 | 0.219 | 0.218 | 0.218 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model $logEBITDA_{i,t+1} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logEBITDA_{i,t+1} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models assess the relationship between a firm's operating performance, measured by EBITDA, in year $t+1$ with Innovation productivity and efficiency variables in year $t$. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of EBITDA and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

### 2.8.6 Internet Appendix F

**Table F1: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.0375 | -0.3850*** | -0.3040*** | -0.3170*** |
| | (6412.6000) | (0.0466) | (0.0469) | (0.0469) |
| RDBE | 0.2910 | 1.0990*** | 0.4260*** | 0.4300*** |
| | (1864.3000) | (0.0853) | (0.0376) | (0.0381) |
| Pat_clean/Book | -0.3010 | | | |
| | (1930.8000) | | | |
| Pat_dirty/Book | 2.7940 | | | |
| | (17910.6000) | | | |
| Pat_emtech/Book | -0.0812 | | | |
| | (520.2000) | | | |
| Pat_other/Book | -0.0397 | | | |
| | (254.8000) | | | |
| Cit/Book | 0.0097 | | | |
| | (61.9500) | | | |
| Cit_clean/Book | | 0.2350 | | |
| | | (0.2060) | | |
| Cit_dirty/Book | | -0.0579 | | |
| | | (0.4590) | | |
| Cit_emtech/Book | | 0.1690 | | |
| | | (0.0916) | | |
| Cit_other/Book | | 0.2510*** | | |
| | | (0.0454) | | |
| Pat/Book | | 0.7520*** | | |
| | | (0.1990) | | |
| Pat_clean/RDC | | | 0.0617 | |
| | | | (0.0591) | |
| Pat_dirty/RDC | | | -0.0369 | |
| | | | (0.0265) | |
| Pat_emtech/RDC | | | 0.3820*** | |
| | | | (0.0747) | |
| Pat_other/RDC | | | 0.0003 | |
| | | | (0.0010) | |
| Cit/RD | | | 0.0110*** | |
| | | | (0.0031) | |
| Cit_clean/RD | | | | 0.0771* |
| | | | | (0.0370) |
| Cit_dirty/RD | | | | -0.0029* |
| | | | | (0.0011) |
| Cit_emtech/RD | | | | 0.1110*** |
| | | | | (0.0242) |
| Cit_other/RD | | | | 0.0068* |
| | | | | (0.0027) |
| Pat/RDC | | | | 0.0021 |
| | | | | (0.0018) |
| EPSlag1-EPSlag6 | 0.8310 | 1.0480*** | 1.1440*** | 1.1530*** |
| | (5328.1000) | (0.1970) | (0.1910) | (0.1930) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 66697 | 66697 | 66696 | 66696 |
| Adjusted $R^2$ | 0.1880 | 0.2120 | 0.1980 | 0.1980 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4(EPSlag1 - EPSlag6) + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \gamma_4(EPSlag1 - EPSlag6) + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. The Innovation productivity and efficiency variables are defined in Table 1 and we refer to EPSlag1 and EPSlag6 as the one year and six year lag of EPS. And we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table F2: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables for firms having non-zero patents during the period 1995-2012**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Intercept | 0.3040*** | 0.3070*** | 0.3070*** | 0.3010*** | 0.3030*** | 0.3030*** |
| | (0.0648) | (0.0645) | (0.0645) | (0.0649) | (0.0659) | (0.0659) |
| RDBE | 0.8380*** | 0.7720*** | 0.7620*** | 0.9800*** | 0.9790*** | 0.9800*** |
| | (0.0848) | (0.0830) | (0.0824) | (0.0926) | (0.0929) | (0.0930) |
| Pat/Book | 0.5590*** | | 0.1450 | | | |
| | (0.1100) | | (0.0922) | | | |
| Cit/Book | | 0.1420*** | 0.1230*** | | | |
| | | (0.0240) | (0.0249) | | | |
| Pat/RDC | | | | 0.0021* | | 0.0002 |
| | | | | (0.0011) | | (0.0005) |
| Cit/RD | | | | | 0.0107*** | 0.0107*** |
| | | | | | (0.0025) | (0.0025) |
| Time FE | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO | NO | NO |
| Observations | 49343 | 49343 | 49343 | 49342 | 49343 | 49342 |
| Adjusted $R^2$ | 0.2270 | 0.2300 | 0.2300 | 0.2220 | 0.2250 | 0.2250 |

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 4-6) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1-3 test whether the knowledge creation process acts as a continuum from R&D to patents to citations. And Models 4-6 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. In the above regression models the sample is the firms having non-zero patents during the period 1995-2012. All the variables are defined in Table 1 and we use the following significance stars $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table F3: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables for firms having non-zero patents during the period 1995-2012**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.3060*** | 0.3070*** | 0.3030*** | 0.3030*** |
| | (0.0645) | (0.0645) | (0.0658) | (0.0658) |
| RDBE | 0.7600*** | 0.7610*** | 0.9800*** | 0.9780*** |
| | (0.0822) | (0.0823) | (0.0929) | (0.0928) |
| Pat_clean/Book | 1.7810** | | | |
| | (0.5670) | | | |
| Pat_dirty/Book | -0.8650 | | | |
| | (0.5120) | | | |
| Pat_other/Book | 0.1050 | | | |
| | (0.0924) | | | |
| Cit/Book | 0.1210*** | | | |
| | (0.0250) | | | |
| Cit_clean/Book | | 0.3190** | | |
| | | (0.1140) | | |
| Cit_dirty/Book | | -0.0718 | | |
| | | (0.0936) | | |
| Cit_other/Book | | 0.1150*** | | |
| | | (0.0260) | | |
| Pat/Book | | 0.1510 | | |
| | | (0.0930) | | |
| Pat_clean/RDC | | | 0.0526 | |
| | | | (0.0351) | |
| Pat_dirty/RDC | | | -0.0291* | |
| | | | (0.0127) | |
| Pat_other/RDC | | | 0.0002 | |
| | | | (0.0005) | |
| Cit/RD | | | 0.0105*** | |
| | | | (0.0024) | |
| Cit_clean/RD | | | | 0.0449* |
| | | | | (0.0213) |
| Cit_dirty/RD | | | | -0.0051 |
| | | | | (0.0043) |
| Cit_other/RD | | | | 0.0097*** |
| | | | | (0.0023) |
| Pat/RDC | | | | 0.0002 |
| | | | | (0.0005) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 49343 | 49343 | 49342 | 49342 |
| Adjusted $R^2$ | 0.2310 | 0.2310 | 0.2260 | 0.2260 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. In the above regression models the sample is the firms having non-zero patents during the period 1995-2012. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table F4: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables for firms having non-zero patents during the period 1995-2012**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.3060*** | 0.3070*** | 0.3030*** | 0.3050*** |
| | (0.0645) | (0.0645) | (0.0657) | (0.0654) |
| RDBE | 0.7610*** | 0.7600*** | 0.9790*** | 0.9670*** |
| | (0.0824) | (0.0823) | (0.0927) | (0.0918) |
| Pat_clean/Book | 1.7710** | | | |
| | (0.5600) | | | |
| Pat_dirty/Book | -0.8490 | | | |
| | (0.5220) | | | |
| Pat_emtech/Book | 0.3900 | | | |
| | (0.3010) | | | |
| Pat_other/Book | 0.0532 | | | |
| | (0.0914) | | | |
| Cit/Book | 0.1190*** | | | |
| | (0.0248) | | | |
| Cit_clean/Book | | 0.3140** | | |
| | | (0.1110) | | |
| Cit_dirty/Book | | -0.0682 | | |
| | | (0.0955) | | |
| Cit_emtech/Book | | 0.1870** | | |
| | | (0.0681) | | |
| Cit_other/Book | | 0.0991*** | | |
| | | (0.0294) | | |
| Pat/Book | | 0.1460 | | |
| | | (0.0924) | | |
| Pat_clean/RDC | | | 0.0413 | |
| | | | (0.0377) | |
| Pat_dirty/RDC | | | -0.0280* | |
| | | | (0.0121) | |
| Pat_emtech/RDC | | | 0.1360*** | |
| | | | (0.0355) | |
| Pat_other/RDC | | | -0.0001 | |
| | | | (0.0003) | |
| Cit/RD | | | 0.0093*** | |
| | | | (0.0023) | |
| Cit_clean/RD | | | | 0.0423* |
| | | | | (0.0207) |
| Cit_dirty/RD | | | | -0.0048 |
| | | | | (0.0043) |
| Cit_emtech/RD | | | | 0.0540*** |
| | | | | (0.0128) |
| Cit_other/RD | | | | 0.0060** |
| | | | | (0.0021) |
| Pat/RDC | | | | 0.0002 |
| | | | | (0.0005) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 49343 | 49343 | 49342 | 49342 |
| Adjusted $R^2$ | 0.2310 | 0.2310 | 0.2270 | 0.2280 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. In the above regression models the sample is the firms having non-zero patents during the period 1995-2012. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table F5: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables for firms which conduct both clean and dirty innovation**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 1.5020*** | 1.5060*** | 1.4820*** | 1.4850*** |
| | (0.0153) | (0.0128) | (0.0153) | (0.0147) |
| RDBE | 0.0055 | 0.0020 | 0.0248 | 0.0230 |
| | (0.0146) | (0.0112) | (0.0146) | (0.0137) |
| Pat_clean/Book | 0.5490* | | | |
| | (0.2770) | | | |
| Pat_dirty/Book | -0.1550 | | | |
| | (0.2020) | | | |
| Pat_emtech/Book | -0.1690* | | | |
| | (0.0778) | | | |
| Pat_other/Book | -0.0219 | | | |
| | 0.0506 | | | |
| Cit/Book | 0.0303** | | | |
| | (0.0112) | | | |
| Cit_clean/Book | | 0.1230*** | | |
| | | (0.0220) | | |
| Cit_dirty/Book | | -0.0005 | | |
| | | (0.0087) | | |
| Cit_emtech/Book | | 0.0106 | | |
| | | (0.0066) | | |
| Cit_other/Book | | 0.0205 | | |
| | | (0.0144) | | |
| Pat/Book | | -0.0148 | | |
| | | (0.0287) | | |
| Pat_clean/RDC | | | -0.0034 | |
| | | | (0.0056) | |
| Pat_dirty/RDC | | | -0.0046 | |
| | | | (0.0029) | |
| Pat_emtech/RDC | | | 0.0339* | |
| | | | (0.0161) | |
| Pat_other/RDC | | | -0.0041 | |
| | | | (0.0021) | |
| Cit/RD | | | 0.0012 | |
| | | | (0.0018) | |
| Cit_clean/RD | | | | 0.0182 |
| | | | | (0.0135) |
| Cit_dirty/RD | | | | -0.0028** |
| | | | | (0.0009) |
| Cit_emtech/RD | | | | 0.0064 |
| | | | | (0.0035) |
| Cit_other/RD | | | | -0.0013 |
| | | | | (0.0006) |
| Pat/RDC | | | | -0.0004 |
| | | | | (0.0008) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 6593 | 6593 | 6593 | 6593 |
| Adjusted $R^2$ | 0.2160 | 0.2180 | 0.1990 | 0.2060 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et. al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. In the above regression models the sample is the firms producing both clean and dirty technologies. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

### 2.8.7 Internet Appendix G: Industry effects

**Table G1: Tobin's Q as a function of disaggregated innovation productivity variables, including emerging technology variants of innovation productivity and interaction between disaggregated innovation productivity variables with the indicator variable, Emtech_firm.**

| | (1) | (2) |
|---|---|---|
| Intercept | 0.195*** | 0.195*** |
| | (0.0392) | (0.0393) |
| RDBE | 1.071*** | 1.071*** |
| | (0.0778) | (0.0778) |
| Pat_clean/Book | 1.478 | |
| | (0.933) | |
| Pat_dirty/Book | -0.993 | |
| | (0.540) | |
| Pat_emtech/Book | 0.634 | |
| | (0.353) | |
| Pat_other/Book | 0.0850 | |
| | (0.108) | |
| $Pat\_clean/Book * Emtech\_firm$ | 0.591 | |
| | (1.214) | |
| $Pat\_dirty/Book * Emtech\_firm$ | 0.594 | |
| | (2.027) | |
| Cit/Book | 0.140*** | |
| | (0.0275) | |
| Cit_clean/Book | | -0.00384 |
| | | (0.0181) |
| Cit_dirty/Book | | -0.0718 |
| | | (0.0870) |
| Cit_emtech/Book | | 0.242** |
| | | (0.0804) |
| Cit_other/Book | | 0.117*** |
| | | (0.0328) |
| $Cit\_clean/Book * Emtech\_firm$ | | 0.533** |
| | | (0.186) |
| $Cit\_dirty/Book * Emtech\_firm$ | | -0.162 |
| | | (0.516) |
| Pat/Book | | 0.201 |
| | | (0.107) |
| Time FE | YES | YES |
| Industry FE | YES | YES |
| Firm-level controls | NO | NO |
| Observations | 79285 | 79285 |
| Adjusted $R^2$ | 0.2160 | 0.2160 |

Notes. The Table presents the regression results of the Models $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat\_clean/Book_{it} + \gamma_3 Pat\_dirty/Book_{it} + \gamma_4 Pat\_other/Book_{it} + \gamma_5 Pat\_emtech/Book_{it} + \gamma_6 Cit/Book_{it} + \mu_1 Pat\_clean/Book_{it} * Emtech\_firm + \mu_2 Pat\_dirty/Book_{it} * Emtech\_firm + \sum_{j=1}^{48} \nu_j industry_j + \sum_{l=1996}^{2012} \kappa_l year_l) + \varepsilon_{it}$ and $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Cit\_clean/Book_{it} + \gamma_3 Cit\_dirty/Book_{it} + \gamma_4 Cit\_other/Book_{it} + \gamma_5 Cit\_emtech/Book_{it} + \gamma_6 Pat/Book_{it} + \mu_1 Cit\_clean/Book_{it} * Emtech\_firm + \mu_2 Cit\_dirty/Book_{it} * Emtech\_firm + \sum_{j=1}^{48} \nu_j industry_j + \sum_{l=1996}^{2012} \kappa_l year_l) + \varepsilon_{it}$ that are estimated using non-linear least squares method. In the above Models, *Emtech_firm* is an indicator variable that take the value 1 if a firm has an emerging technology patent published by the USPTO and 0 otherwise. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity and Cit/Book as a proxy for citation productivity. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table G2: Tobin's Q as a function of disaggregated Innovation patent productivity variables, industry sectors and the interaction between the patent productivity variables and Drugs industry sector**

| | (1) | (2) | (3) |
|---|---|---|---|
| Intercept | 0.728*** | 0.1950*** | 1.273 |
| | (0.113) | (0.0393) | (6524.6) |
| RDBE | 0.846*** | 1.0720*** | 0.348 |
| | (0.151) | (0.0778) | (2312.5) |
| Pat_clean/Book | -6.869** | 1.8030** | 0.612 |
| | (2.245) | (0.6150) | (0.6150) |
| Pat_dirty/Book | 213.6 | -0.9720 | -0.329 |
| | (129.2) | (0.5520) | (2148.0) |
| Pat_other/Book | 0.124 | 0.1700 | 0.0519 |
| | (0.280) | (0.1090) | (348.4) |
| Cit/Book | 0.0786* | 0.1440*** | 0.0479 |
| | (0.0356) | (0.0277) | (333.0) |
| Industry = Drugs | | | -0.505 |
| | | | (3460.3) |
| Pat_clean/Book*Drugs | | | -3.742 |
| | | | (24831.0) |
| Pat_dirty/Book*Drugs | | | 125.3 |
| | | | (791424.4) |
| Time FE | YES | YES | YES |
| Industry FE | NO | YES | YES |
| Firm-level controls | NO | NO | NO |
| Observations | 3767 | 79285 | 79285 |
| Adjusted $R^2$ | 0.155 | 0.2150 | 0.214 |

Notes. The Table presents the regression results of various specifications of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat\_clean/Book_{it} + \gamma_3 Pat\_dirty/Book_{it} + \gamma_3 Pat\_other/Book_{it} + \gamma_4 Cit/Book_{it} + \mu_1 Pat\_clean/Book_{it} * Drugs + \mu_2 Pat\_dirty/Book_{it} * Drugs + \sum_{j=1}^{48} \nu_j industry_j + \sum_{l=1996}^{2012} \kappa_l year_l) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. We choose the Pharmaceutical Products (henceforth Drugs) industry and focus on its interaction with our clean and dirty innovation patent productivity variables. The sample of firms in the first regression model (column 1) of the Table belong to the Drugs industry sector. In the other regression models (columns 2 and 3) of the Table the sample of firms is the whole sample of firms in our data set. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table G3: Tobin's Q as a function of disaggregated Innovation citation productivity variables, industry sectors and the interaction between the citation productivity variables and Drugs industry sector**

| | (1) | (2) | (3) |
|---|---|---|---|
| Intercept | 0.728*** | 0.1950*** | 2.587 |
| | (0.112) | (0.0393) | (2377.1) |
| RDBE | 0.840*** | 1.0720*** | 0.0977 |
| | (0.122) | (0.0779) | (229.0) |
| Cit_clean/Book | -5.237*** | 0.3220** | 0.0293 |
| | (1.091) | (0.1170) | (69.03) |
| Cit_dirty/Book | 144.4* | -0.0876 | -0.00813 |
| | (61.04) | (0.1050) | (19.49) |
| Cit_other/Book | 0.0788* | 0.1390*** | 0.0127 |
| | (0.0307) | (0.0291) | (30.02) |
| Pat/Book | 0.134 | 0.2160* | 0.0198 |
| | (0.175) | (0.1080) | (46.06) |
| Industry = Drugs | | | -0.864 |
| | | | (305.8) |
| Cit_clean/Book*Drugs | | | -0.867 |
| | | | (2074.7) |
| Cit_dirty/Book*Drugs | | | 23.63 |
| | | | (54767.6) |
| Time FE | YES | YES | YES |
| Industry FE | NO | YES | YES |
| Firm-level controls | NO | NO | NO |
| Observations | 3767 | 79285 | 79285 |
| Adjusted $R^2$ | 0.154 | 0.2150 | 0.216 |

Notes. The Table presents the regression results of various specifications of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Cit\_clean/Book_{it} + \gamma_3 Cit\_dirty/Book_{it} + \gamma_3 Cit\_other/Book_{it} + \gamma_4 Pat/Book_{it} + \mu_1 Cit\_clean/Book_{it} * Drugs + \mu_2 Cit\_dirty/Book_{it} * Drugs + \sum_{j=1}^{48} \nu_j industry_j + \sum_{l=1996}^{2012} \kappa_l year_l) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. We choose the Pharmaceutical Products (henceforth Drugs) industry and focus on its interaction with our clean and dirty innovation citation productivity variables. The sample of firms in the first regression model (column 1) of the Table belong to the Drugs industry sector. In the other regression models (columns 2 and 3) of the Table the sample of firms is the whole sample of firms in our data set. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

### 2.8.8 Internet Appendix H: Country Fixed effects

**Table H1: Tobin's Q as a function of aggregated Innovation productivity and efficiency variables and controlling for country fixed effects**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Intercept | 0.486*** | 0.484*** | 0.484*** | 0.487*** | 0.484*** | 0.484*** |
| | (0.0460) | (0.0461) | (0.0460) | (0.0464) | (0.0462) | (0.0462) |
| RDBE | 0.490*** | 0.480*** | 0.472*** | 0.549*** | 0.549*** | 0.549*** |
| | (0.0451) | (0.0455) | (0.0453) | (0.0465) | (0.0464) | (0.0464) |
| Pat/Book | 0.337*** | | 0.155* | | | |
| | (0.0775) | | (0.0773) | | | |
| Cit/Book | | 0.0732*** | 0.0531** | | | |
| | | (0.0162) | (0.0176) | | | |
| Pat/RDC | | | | 0.00232* | | 0.000568 |
| | | | | (0.000983) | | (0.000532) |
| Cit/RD | | | | | 0.00732*** | 0.00721*** |
| | | | | | (0.00162) | (0.00163) |
| Time FE | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO | NO | NO |
| Observations | 79285 | 79285 | 79285 | 79284 | 79285 | 79284 |
| Adjusted $R^2$ | 0.309 | 0.309 | 0.309 | 0.307 | 0.309 | 0.309 |

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/Book_{it} + \gamma_3 Cit/Book_{it} + \sum_{i=1996}^{2012} \kappa_i year_i + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 4-6) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{i=2}^{17} \kappa_i year_i + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1-3 test whether the knowledge creation process acts as a continuum from R&D to patents to citations. And Models 4-6 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table H2: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables and controlling for country fixed effects**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.483*** | 0.484*** | 0.484*** | 0.484*** |
| | (0.0461) | (0.0460) | (0.0462) | (0.0462) |
| RDBE | 0.472*** | 0.472*** | 0.550*** | 0.549*** |
| | (0.0453) | (0.0454) | (0.0464) | (0.0464) |
| Pat_clean/Book | 1.138* | | | |
| | (0.449) | | | |
| Pat_dirty/Book | -0.550 | | | |
| | (0.350) | | | |
| Pat_other/Book | 0.134 | | | |
| | (0.0782) | | | |
| Cit/Book | 0.0524** | | | |
| | (0.0178) | | | |
| Cit_clean/Book | | 0.101 | | |
| | | (0.0561) | | |
| Cit_dirty/Book | | -0.0661 | | |
| | | (0.0781) | | |
| Cit_other/Book | | 0.0507** | | |
| | | (0.0187) | | |
| Pat/Book | | 0.164* | | |
| | | (0.0784) | | |
| Pat_clean/RDC | | | 0.0251 | |
| | | | (0.0178) | |
| Pat_dirty/RDC | | | -0.0157 | |
| | | | (0.00899) | |
| Pat_other/RDC | | | 0.000510 | |
| | | | (0.000517) | |
| Cit/RD | | | 0.00717*** | |
| | | | (0.00161) | |
| Cit_clean/RD | | | | 0.0184 |
| | | | | (0.0131) |
| Cit_dirty/RD | | | | -0.00258 |
| | | | | (0.00341) |
| Cit_other/RD | | | | 0.00703*** |
| | | | | (0.00157) |
| Pat/RDC | | | | 0.000559 |
| | | | | (0.000527) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79285 | | 79284 | 79284 |
| Adjusted $R^2$ | 0.309 | 0.309 | 0.309 | 0.309 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table H3: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables and controlling for country fixed effects, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.328*** | 0.328*** | 0.324*** | 0.324*** |
| | (0.0206) | (0.0208) | (0.0200) | (0.0207) |
| RDBE | 0.182*** | 0.181*** | 0.194*** | 0.194*** |
| | (0.0185) | (0.0185) | (0.0247) | (0.0247) |
| Pat_clean/Book | 1.739** | | | |
| | (0.777) | | | |
| Pat_dirty/Book | -1.352 | | | |
| | (1.193) | | | |
| Pat_other/Book | 0.0835 | | | |
| | (0.0618) | | | |
| Cit/Book | 0.0154 | | | |
| | (0.0124) | | | |
| Cit_clean/Book | | 0.167* | | |
| | | (0.0850) | | |
| Cit_dirty/Book | | -0.458* | | |
| | | (0.236) | | |
| Cit_other/Book | | 0.0159 | | |
| | | (0.0121) | | |
| Pat/Book | | 0.0842 | | |
| | | (0.0571) | | |
| Pat_clean/RDC | | | 0.0899** | |
| | | | (0.0387) | |
| Pat_dirty/RDC | | | 0.0200 | |
| | | | (0.0201) | |
| Pat_other/RDC | | | 0.00193** | |
| | | | (0.000888) | |
| Cit/RD | | | 0.00296*** | |
| | | | (0.000790) | |
| Cit_clean/RD | | | | 0.0154** |
| | | | | (0.00662) |
| Cit_dirty/RD | | | | -0.00497 |
| | | | | (0.00740) |
| Cit_other/RD | | | | 0.00328*** |
| | | | | (0.000915) |
| Pat/RDC | | | | 0.00245* |
| | | | | (0.00121) |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79,285 | 79,285 | 79,284 | 79,284 |
| avg. R-squared | 0.329 | 0.328 | 0.327 | 0.327 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table H4: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables and controlling for country fixed effects for firms which conduct both clean and dirty innovation**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 1.408*** | 1.405*** | 1.410*** | 1.409*** |
| | (0.0254) | (0.0244) | (0.0274) | (0.0268) |
| RDBE | -0.0104 | -0.0200 | 0.00230 | 0.00192 |
| | (0.0120) | (0.0104) | (0.00730) | (0.00710) |
| Pat_clean/Book | 0.571** | | | |
| | (0.208) | | | |
| Pat_dirty/Book | -0.173 | | | |
| | (0.121) | | | |
| Pat_other/Book | -0.000643 | | | |
| | (0.0403) | | | |
| Cit/Book | 0.0177 | | | |
| | (0.0101) | | | |
| Cit_clean/Book | | 0.0900*** | | |
| | | (0.0200) | | |
| Cit_dirty/Book | | -0.0135 | | |
| | | (0.0275) | | |
| Cit_other/Book | | 0.00436 | | |
| | | (0.00829) | | |
| Pat/Book | | 0.0419 | | |
| | | (0.0330) | | |
| Pat_clean/RDC | | | -0.00198 | |
| | | | (0.00274) | |
| Pat_dirty/RDC | | | -0.00384* | |
| | | | (0.00170) | |
| Pat_other/RDC | | | 0.00112 | |
| | | | (0.00193) | |
| Cit/RD | | | 0.00115 | |
| | | | (0.00129) | |
| Cit_clean/RD | | | | 0.00683 |
| | | | | (0.00875) |
| Cit_dirty/RD | | | | -0.00275 |
| | | | | (0.00168) |
| Cit_other/RD | | | | 0.00100 |
| | | | | (0.00109) |
| Pat/RDC | | | | -0.000365 |
| | | | | (0.000680) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 6198 | 6198 | 6198 | 6198 |
| Adjusted $R^2$ | 0.350 | 0.352 | 0.341 | 0.342 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. In the above regression models the sample is the firms producing both clean and dirty technologies. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table H5: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables and controlling for country fixed effects**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.483*** | 0.483*** | 0.483*** | 0.483*** |
| | (0.0460) | (0.0460) | (0.0459) | (0.0460) |
| RDBE | 0.472*** | 0.472*** | 0.546*** | 0.545*** |
| | (0.0454) | (0.0453) | (0.0462) | (0.0461) |
| Pat_clean/Book | 1.137* | | | |
| | (0.445) | | | |
| Pat_dirty/Book | -0.543 | | | |
| | (0.351) | | | |
| Pat_emtech/Book | 0.352 | | | |
| | (0.233) | | | |
| Pat_other/Book | 0.0921 | | | |
| | (0.0787) | | | |
| Cit/Book | 0.0517** | | | |
| | (0.0178) | | | |
| Cit_clean/Book | | 0.101 | | |
| | | (0.0557) | | |
| Cit_dirty/Book | | -0.0650 | | |
| | | (0.0783) | | |
| Cit_emtech/Book | | 0.103* | | |
| | | (0.0480) | | |
| Cit_other/Book | | 0.0385 | | |
| | | (0.0216) | | |
| Pat/Book | | 0.163* | | |
| | | (0.0780) | | |
| Pat_clean/RDC | | | 0.0116 | |
| | | | (0.0174) | |
| Pat_dirty/RDC | | | -0.0133 | |
| | | | (0.00846) | |
| Pat_emtech/RDC | | | 0.123*** | |
| | | | (0.0272) | |
| Pat_other/RDC | | | 0.0000691 | |
| | | | (0.000301) | |
| Cit/RD | | | 0.00615*** | |
| | | | (0.00153) | |
| Cit_clean/RD | | | | 0.0173 |
| | | | | (0.0125) |
| Cit_dirty/RD | | | | -0.00226 |
| | | | | (0.00337) |
| Cit_emtech/RD | | | | 0.0380*** |
| | | | | (0.00852) |
| Cit_other/RD | | | | 0.00424** |
| | | | | (0.00138) |
| Pat/RDC | | | | 0.000565 |
| | | | | (0.000549) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79285 | 79285 | 79284 | 79284 |
| Adjusted $R^2$ | 0.310 | 0.309 | 0.310 | 0.310 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et. al., 2005. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

**Table H6: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables, and controlling for country fixed effects, esestimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.327*** | 0.327*** | 0.323*** | 0.324*** |
| | (0.0207) | (0.0208) | (0.0200) | (0.0207) |
| RDBE | 0.181*** | 0.181*** | 0.194*** | 0.194*** |
| | (0.0182) | (0.0181) | (0.0247) | (0.0247) |
| Pat_clean/Book | 1.731** | | | |
| | (0.769) | | | |
| Pat_dirty/Book | -1.370 | | | |
| | (1.196) | | | |
| Pat_emtech/Book | 0.238* | | | |
| | (0.120) | | | |
| Pat_other/Book | 0.0701 | | | |
| | (0.0572) | | | |
| Cit/Book | 0.0175 | | | |
| | (0.0118) | | | |
| Cit_clean/Book | | 0.168* | | |
| | | (0.0832) | | |
| Cit_dirty/Book | | -0.460* | | |
| | | (0.238) | | |
| Cit_emtech/Book | | 0.118** | | |
| | | (0.0410) | | |
| Cit_other/Book | | 0.0125 | | |
| | | (0.0121) | | |
| Pat/Book | | 0.0603 | | |
| | | (0.0555) | | |
| Pat_clean/RDC | | | 0.0762* | |
| | | | (0.0410) | |
| Pat_dirty/RDC | | | 0.0267 | |
| | | | (0.0189) | |
| Pat_emtech/RDC | | | 0.0726*** | |
| | | | (0.0165) | |
| Pat_other/RDC | | | -0.000279 | |
| | | | (0.000872) | |
| Cit/RD | | | 0.00261*** | |
| | | | (0.000787) | |
| Cit_clean/RD | | | | 0.0167** |
| | | | | (0.00593) |
| Cit_dirty/RD | | | | -0.00486 |
| | | | | (0.00750) |
| Cit_emtech/RD | | | | 0.0202*** |
| | | | | (0.00648) |
| Cit_other/RD | | | | 0.00198** |
| | | | | 0.000706 |
| Pat/RDC | | | | 0.00242* |
| | | | | (0.00124) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79,285 | 79,285 | 79,284 | 79,284 |
| avg. R-squared | 0.329 | 0.329 | 0.328 | 0.328 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table H7: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables and controlling for firm traits and country fixed effects, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.328*** | 0.329*** | 0.322*** | 0.322*** |
| | (0.0226) | (0.0225) | (0.0228) | (0.0226) |
| RDBE | 0.946*** | 0.951*** | 1.035*** | 1.036*** |
| | (0.0654) | (0.0656) | (0.0659) | (0.0659) |
| Pat_clean/Book | 2.229** | | | |
| | (1.007) | | | |
| Pat_dirty/Book | -0.733 | | | |
| | (1.279) | | | |
| Pat_other/Book | 0.0917 | | | |
| | (0.120) | | | |
| Cit/Book | 0.0946*** | | | |
| | (0.0205) | | | |
| Cit_clean/Book | | 0.435** | | |
| | | (0.190) | | |
| Cit_dirty/Book | | -0.230 | | |
| | | (0.266) | | |
| Cit_other/Book | | 0.0884*** | | |
| | | (0.0199) | | |
| Pat/Book | | 0.105 | | |
| | | (0.118) | | |
| Pat_clean/RDC | | | 0.156*** | |
| | | | (0.0510) | |
| Pat_dirty/RDC | | | 0.0333 | |
| | | | (0.0324) | |
| Pat_other/RDC | | | 0.00167 | |
| | | | (0.00152) | |
| Cit/RD | | | 0.00403** | |
| | | | (0.00143) | |
| Cit_clean/RD | | | | 0.0379** |
| | | | | (0.0153) |
| Cit_dirty/RD | | | | 0.0149 |
| | | | | (0.00984) |
| Cit_other/RD | | | | 0.00462*** |
| | | | | (0.00157) |
| Pat/RDC | | | | 0.00169 |
| | | | | (0.00163) |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | YES | YES | YES | YES |
| Observations | 50,494 | 50,494 | 50,493 | 50,493 |
| avg. R-squared | 0.449 | 0.449 | 0.448 | 0.448 |

Notes. The Table presents the regression results of various specifications (columns 1, 2, 5 and 6) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earningabnormalit + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3, 4, 7 and 8) $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earningabnormalit + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005 and Hirshleifer et al., 2013 with the inclusion of firm-level control variables, year and industry fixed-effects. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process,from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

### 2.8.9 Internet Appendix I: Reconstructed Innovation productivity variables

**Table I1: Tobin's Q as a function of reconstructed aggregated Innovation productivity variables and aggregated Innovation efficiency variables**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Intercept | 0.430*** | 0.426*** | 0.427*** | 0.429*** | 0.427*** | 0.427*** |
| | (0.0450) | (0.0451) | (0.0451) | (0.0454) | (0.0452) | (0.0452) |
| RDTA | 2.217*** | 2.202*** | 2.163*** | 2.469*** | 2.466*** | 2.466*** |
| | (0.157) | (0.157) | (0.157) | (0.163) | (0.162) | (0.162) |
| Pat/Total_assets | 1.107*** | | 0.519*** | | | |
| | (0.167) | | (0.155) | | | |
| Cit/Total_assets | | 0.234*** | 0.166*** | | | |
| | | (0.0354) | (0.0348) | | | |
| Pat/RDC | | | | 0.00246* | | 0.000710 |
| | | | | (0.00100) | | (0.000578) |
| Cit/RD | | | | | 0.00682*** | 0.00668*** |
| | | | | | (0.00158) | (0.00159) |
| Time FE | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO | NO | NO |
| Observations | 79285 | 79285 | 79285 | 79284 | 79285 | 79284 |
| Adjusted $R^2$ | 0.338 | 0.339 | 0.339 | 0.334 | 0.335 | 0.335 |

Notes. The Table presents the regression results of various specifications (columns 1-3) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDTA_{it} + \gamma_2 Pat/Total\_assets_{it} + \gamma_3 Cit/Total\_assets_{it} + \sum_{i=1996}^{2012} \kappa_i year_i + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 4-6), $logQ_{it} = \alpha + log(1 + \gamma_1 RDTA_{it} + \gamma_2 Pat/RDC_{it} + \gamma_3 Cit/RD_{it} + \sum_{i=2}^{17} \kappa_i year_i + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. We define RDTA as R&D expense over book value of total assets and use this as a proxy for R&D productivity. Similarly, we reconstruct the patent and citation productivity variables by employing book value of total assets as denominator instead of book value of equity. Hence our reconstructed proxies for patent and citation productivity variables, Pat/Total_assets and Cit/Total_assets, are defined as number of US patents of the firm, in any patent category, divided by book value of total assets and adjusted patent citation of a firm divided by book value of total assets, respectively. Models 1-3 test whether the knowledge creation process acts as a continuum from R&D to patents to citations. And Models 4-6 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating patents and citations. In our specifications we use Pat/RDC as a proxy for patent efficiency and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table I2: Tobin's Q as a function of reconstructed disaggregated Innovation productivity variables and disaggregated Innovation efficiency variables**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.426*** | 0.426*** | 0.426*** | 0.426*** |
| | (0.0451) | (0.0451) | (0.0452) | (0.0452) |
| RDTA | 2.165*** | 2.166*** | 2.467*** | 2.466*** |
| | (0.157) | (0.157) | (0.162) | (0.162) |
| Pat_clean/Total_assets | 2.307* | | | |
| | (0.919) | | | |
| Pat_dirty/Total_assets | -1.688 | | | |
| | (1.835) | | | |
| Pat_other/Total_assets | 0.479** | | | |
| | (0.159) | | | |
| Cit/Total_assets | 0.164*** | | | |
| | (0.0351) | | | |
| Cit_clean/Total_assets | | 0.337* | | |
| | | (0.136) | | |
| Cit_dirty/Total_assets | | -0.616*** | | |
| | | (0.118) | | |
| Cit_other/Total_assets | | 0.162*** | | |
| | | (0.0370) | | |
| Pat/Total_assets | | 0.516*** | | |
| | | (0.155) | | |
| Pat_clean/RDC | | | 0.0233 | |
| | | | (0.0166) | |
| Pat_dirty/RDC | | | -0.0155 | |
| | | | (0.00936) | |
| Pat_other/RDC | | | 0.000648 | |
| | | | (0.000565) | |
| Cit/RD | | | 0.00665*** | |
| | | | (0.00158) | |
| Cit_clean/RD | | | | 0.0162 |
| | | | | (0.0127) |
| Cit_dirty/RD | | | | -0.00232 |
| | | | | (0.00360) |
| Cit_other/RD | | | | 0.00656*** |
| | | | | (0.00155) |
| Pat/RDC | | | | 0.000702 |
| | | | | (0.000574) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79285 | 79285 | 79284 | 79284 |
| Adjusted $R^2$ | 0.339 | 0.339 | 0.335 | 0.335 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDTA_{it} + \gamma_2 Pat^*/Total\_assets_{it} + \gamma_3 Cit^*/Total\_assets_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDTA_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. We define RDTA as R&D expense over book value of total assets and use this as a proxy for R&D productivity. Similarly, we reconstruct the patent and citation productivity variables by employing book value of total assets as denominator instead of book value of equity. Hence our reconstructed proxies for patent and citation productivity variables, Pat/Total_assets and Cit/Total_assets, are defined as number of US patents of the firm, in any patent category, divided by book value of total assets and adjusted patent citation of a firm divided by book value of total assets, respectively. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use Pat/RDC as a proxy for patent efficiency and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table I3: Tobin's Q as a function of reconstructed disaggregated Innovation productivity variables and disaggregated Innovation efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.269*** | 0.268*** | 0.264*** | 0.265*** |
| | (0.0168) | (0.0168) | (0.0168) | (0.0174) |
| RDTA | 1.418*** | 1.415*** | 1.531*** | 1.531*** |
| | (0.127) | (0.127) | (0.114) | (0.115) |
| Pat_clean/Total_assets | 3.027** | | | |
| | (1.056) | | | |
| Pat_dirty/Total_assets | -3.220 | | | |
| | (2.534) | | | |
| Pat_other/Total_assets | 0.229** | | | |
| | (0.0870) | | | |
| Cit/Total_assets | 0.0394 | | | |
| | (0.0248) | | | |
| Cit_clean/Total_assets | | 0.380** | | |
| | | (0.172) | | |
| Cit_dirty/Total_assets | | -1.535** | | |
| | | (0.675) | | |
| Cit_other/Total_assets | | 0.0404 | | |
| | | (0.0247) | | |
| Pat/Total_assets | | 0.239** | | |
| | | (0.0876) | | |
| Pat_clean/RDC | | | 0.0836** | |
| | | | (0.0377) | |
| Pat_dirty/RDC | | | 0.0272 | |
| | | | (0.0215) | |
| Pat_other/RDC | | | 0.00220** | |
| | | | (0.000900) | |
| Cit/RD | | | 0.00276*** | |
| | | | (0.000718) | |
| Cit_clean/RD | | | | 0.0145** |
| | | | | (0.00642) |
| Cit_dirty/RD | | | | -0.00477 |
| | | | | (0.00731) |
| Cit_other/RD | | | | 0.00307*** |
| | | | | (0.000864) |
| Pat/RDC | | | | 0.00271** |
| | | | | (0.00119) |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79,285 | 79,285 | 79,284 | 79,284 |
| avg. R-squared | 0.358 | 0.358 | 0.354 | 0.354 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model $logQ_{it} = \alpha + \gamma_1 RDTA_{it} + \gamma_2 Pat^*/Total\_assets_{it} + \gamma_3 Cit^*/Total\_assets_{it} + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logQ_{it} = \alpha + \gamma_1 RDTA_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. We define RDTA as R&D expense over book value of total assets and use this as a proxy for R&D productivity. Similarly, we reconstruct the patent and citation productivity variables by employing book value of total assets as denominator instead of book value of equity. Hence our reconstructed proxies for patent and citation productivity variables, Pat/Total_assets and Cit/Total_assets, are defined as number of US patents of the firm, in any patent category, divided by book value of total assets and adjusted patent citation of a firm divided by book value of total assets, respectively. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use Pat/RDC as a proxy for patent efficiency and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table I4: Tobin's Q as a function of reconstructed disaggregated Innovation productivity variables and disaggregated Innovation efficiency variables and controlling for firm traits, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.296*** | 0.296*** | 0.290*** | 0.291*** |
| | (0.0210) | (0.0206) | (0.0215) | (0.0213) |
| RDTA | 2.968*** | 2.971*** | 3.161*** | 3.162*** |
| | (0.264) | (0.265) | (0.250) | (0.249) |
| Pat_clean/Total_assets | 5.486** | | | |
| | (2.514) | | | |
| Pat_dirty/Total_assets | 2.074 | | | |
| | (3.303) | | | |
| Pat_other/Total_assets | 0.0319 | | | |
| | (0.216) | | | |
| Cit/Total_assets | 0.191*** | | | |
| | (0.0367) | | | |
| Cit_clean/Total_assets | | 1.003*** | | |
| | | (0.290) | | |
| Cit_dirty/Total_assets | | 0.897 | | |
| | | (0.871) | | |
| Cit_other/Total_assets | | 0.187*** | | |
| | | (0.0368) | | |
| Pat/Total_assets | | 0.0404 | | |
| | | (0.212) | | |
| Pat_clean/RDC | | | 0.144*** | |
| | | | (0.0477) | |
| Pat_dirty/RDC | | | 0.0327 | |
| | | | (0.0329) | |
| Pat_other/RDC | | | 0.00196 | |
| | | | (0.00150) | |
| Cit/RD | | | 0.00382** | |
| | | | (0.00134) | |
| Cit_clean/RD | | | | 0.0342** |
| | | | | (0.0134) |
| Cit_dirty/RD | | | | 0.0154 |
| | | | | (0.0105) |
| Cit_other/RD | | | | 0.00440*** |
| | | | | (0.00148) |
| Pat/RDC | | | | 0.00194 |
| | | | | (0.00160) |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | YES | YES | YES | YES |
| Observations | 50,494 | 50,494 | 50,493 | 50,493 |
| avg. R-squared | 0.467 | 0.467 | 0.465 | 0.465 |

Notes. The Table presents the regression results of various specifications (columns 1, 2, 5 and 6) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDTA_{it} + \gamma_2 Pat^*/Total\_assets_{it} + \gamma_3 Cit^*/Total\_assets_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earningabnormalit + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3, 4, 7 and 8) $logQ_{it} = \alpha + log(1 + \gamma_1 RDTA_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 \gamma_3 Cit^*/RD_{it} + \gamma_4 RDG_{it} + \gamma_5 invBE_{it} + \gamma_6 taxRDBE_{it} + \gamma_7 CEME_{it} + \gamma_8 Earningabnormalit + \gamma_9 Adverts_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005 and Hirshleifer et al., 2013 with the inclusion of firm-level control variables, country and industry fixed-effects. We define RDTA as R&D expense over book value of total assets and use this as a proxy for R&D productivity. Similarly, we reconstruct the patent and citation productivity variables by employing book value of total assets as denominator instead of book value of equity. Hence our reconstructed proxies for patent and citation productivity variables, Pat/Total_assets and Cit/Total_assets, are defined as number of US patents of the firm, in any patent category, divided by book value of total assets and adjusted patent citation of a firm divided by book value of total assets, respectively. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process,from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table I5: Tobin's Q as a function of reconstructed disaggregated Innovation productivity variables and disaggregated Innovation efficiency variables for firms which conduct both clean and dirty innovation**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 1.471*** | 1.399*** | 1.405*** | 1.405*** |
| | (0.0236) | (0.0240) | (0.0233) | (0.0233) |
| RDTA | 0.551*** | 0.594*** | 0.717*** | 0.702*** |
| | (0.151) | (0.161) | (0.156) | (0.156) |
| Pat_clean/Total_assets | 0.925** | | | |
| | (0.305) | | | |
| Pat_dirty/Total_assets | -0.389 | | | |
| | (0.705) | | | |
| Pat_other/Total_assets | 0.0884 | | | |
| | (0.101) | | | |
| Cit/Total_assets | 0.0192 | | | |
| | (0.0144) | | | |
| Cit_clean/Total_assets | | 0.0876*** | | |
| | | (0.0224) | | |
| Cit_dirty/Total_assets | | -0.145 | | |
| | | (0.143) | | |
| Cit_other/Total_assets | | -0.00966 | | |
| | | (0.0146) | | |
| Pat/Total_assets | | 0.239* | | |
| | | (0.112) | | |
| Pat_clean/RDC | | | -0.00243 | |
| | | | (0.00260) | |
| Pat_dirty/RDC | | | -0.00347 | |
| | | | (0.00177) | |
| Pat_other/RDC | | | 0.00163 | |
| | | | (0.00180) | |
| Cit/RD | | | 0.000956 | |
| | | | (0.00109) | |
| Cit_clean/RD | | | | 0.00530 |
| | | | | (0.00823) |
| Cit_dirty/RD | | | | -0.00230 |
| | | | | (0.00174) |
| Cit_other/RD | | | | 0.00107 |
| | | | | (0.00104) |
| Pat/RDC | | | | -0.0000736 |
| | | | | (0.000711) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 6198 | 6198 | 6198 | 6198 |
| Adjusted $R^2$ | 0.367 | 0.368 | 0.359 | 0.360 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDTA_{it} + \gamma_2 Pat^*/Total\_assets_{it} + \gamma_3 Cit^*/Total\_assets_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4) $logQ_{it} = \alpha + log(1 + \gamma_1 RDTA_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. We define RDTA as R&D expense over book value of total assets and use this as a proxy for R&D productivity. Similarly, we reconstruct the patent and citation productivity variables by employing book value of total assets as denominator instead of book value of equity. Hence our reconstructed proxies for patent and citation productivity variables, Pat/Total_assets and Cit/Total_assets, are defined as number of US patents of the firm, in any patent category, divided by book value of total assets and adjusted patent citation of a firm divided by book value of total assets, respectively. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use Pat/RDC as a proxy for patent efficiency and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. In the above regression models the sample is the firms producing both clean and dirty technologies. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table I6: Tobin's Q as a function of reconstructed disaggregated Innovation productivity variables and disaggregated Innovation efficiency variables, including emerging technology variants of Innovation productivity and efficiency variables**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.426*** | 0.426*** | 0.426*** | 0.426*** |
| | (0.0450) | (0.0451) | (0.0450) | (0.0450) |
| RDTA | 2.161*** | 2.163*** | 2.454*** | 2.454*** |
| | (0.157) | (0.157) | (0.161) | (0.162) |
| Pat_clean/Total_assets | 2.287* | | | |
| | (0.898) | | | |
| Pat_dirty/Total_assets | -1.547 | | | |
| | (1.870) | | | |
| Pat_other/Total_assets | 0.302 | | | |
| | (0.175) | | | |
| Pat_emtech/Total_assets | 1.256** | | | |
| | (0.461) | | | |
| Cit/Total_assets | 0.162*** | | | |
| | (0.0350) | | | |
| Cit_clean/Total_assets | | 0.340* | | |
| | | (0.133) | | |
| Cit_dirty/Total_assets | | -0.683*** | | |
| | | (0.186) | | |
| Cit_other/Total_assets | | 0.141*** | | |
| | | (0.0424) | | |
| Cit_emtech/Total_assets | | 0.259* | | |
| | | (0.109) | | |
| Pat/Total_assets | | 0.506** | | |
| | | (0.155) | | |
| Pat_clean/RDC | | | 0.00984 | |
| | | | (0.0164) | |
| Pat_dirty/RDC | | | -0.0128 | |
| | | | (0.00864) | |
| Pat_other/RDC | | | 0.000178 | |
| | | | (0.000351) | |
| Pat_emtech/RDC | | | 0.103*** | |
| | | | (0.0243) | |
| Cit/RD | | | 0.00584*** | |
| | | | (0.00150) | |
| Cit_clean/RD | | | | 0.0154 |
| | | | | (0.0121) |
| Cit_dirty/RD | | | | -0.00203 |
| | | | | (0.00356) |
| Cit_other/RD | | | | 0.00414** |
| | | | | (0.00139) |
| Cit_emtech/RD | | | | 0.0317*** |
| | | | | (0.00755) |
| Pat/RDC | | | | 0.000713 |
| | | | | (0.000596) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Country FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 79285 | 79285 | 79284 | 79284 |
| Adjusted $R^2$ | 0.339 | 0.339 | 0.336 | 0.336 |

Notes. The Table presents the regression results of various specifications (columns 1-2) of the Model, $logQ_{it} = \alpha + log(1 + \gamma_1 RDTA_{it} + \gamma_2 Pat^*/Total\_assets_{it} + \gamma_3 Cit^*/Total\_assets_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ and the Model (columns 3-4), $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{l=1996}^{2012} \kappa_l year_l + \sum_{k=2}^{12} \pi_j Country_j + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et. al., 2005. We define RDTA as R&D expense over book value of total assets and use this as a proxy for R&D productivity. Similarly, we reconstruct the patent and citation productivity variables by employing book value of total assets as denominator instead of book value of equity. Hence our reconstructed proxies for patent and citation productivity variables, Pat/Total_assets and Cit/Total_assets, are defined as number of US patents of the firm, in any patent category, divided by book value of total assets and adjusted patent citation of a firm divided by book value of total assets, respectively. Models 1 and 2 test whether the knowledge creation process acts as a continuum from R&D to clean patents to clean citations. And Models 3 and 4 test the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use Pat/RDC as a proxy for patent efficiency and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars $^*$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

### 2.8.10 Internet Appendix J: Grey technologies

**Table J1: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.0380 | 0.0381 | 0.0362 | 0.0361 |
| | (0.0249) | (0.0249) | (0.0254) | (0.0255) |
| RDBE | 0.0899*** | 0.0874*** | 0.0878*** | 0.0877*** |
| | (0.0171) | (0.0173) | (0.0189) | (0.0189) |
| Pat_clean/Book | 1.930* | | | |
| | (0.961) | | | |
| Pat_dirty/Book | -0.429 | | | |
| | (1.349) | | | |
| Pat_other/Book | 0.148* | | | |
| | (0.0731) | | | |
| Pat_grey/Book | 2.213 | | | |
| | (1.407) | | | |
| Cit/Book | 0.0128 | | | |
| | (0.0103) | | | |
| Cit_clean/Book | | 0.348** | | |
| | | (0.133) | | |
| Cit_dirty/Book | | -0.132 | | |
| | | (0.215) | | |
| Cit_other/Book | | 0.0122 | | |
| | | (0.0101) | | |
| Cit_grey/Book | | 0.331* | | |
| | | (0.181) | | |
| Pat/Book | | 0.149** | | |
| | | (0.0694) | | |
| Pat_clean/RDC | | | 0.0973** | |
| | | | (0.0436) | |
| Pat_dirty/RDC | | | -0.0140 | |
| | | | (0.0506) | |
| Pat_other/RDC | | | 0.00313** | |
| | | | (0.00116) | |
| Pat_grey/RDC | | | 0.321* | |
| | | | (0.166) | |
| Cit/RD | | | 0.00168*** | |
| | | | (0.000431) | |
| Cit_clean/RD | | | | 0.0287** |
| | | | | (0.0115) |
| Cit_dirty/RD | | | | 0.000390 |
| | | | | (0.0111) |
| Cit_other/RD | | | | 0.00263** |
| | | | | (0.000924) |
| Cit_grey/RD | | | | 0.0529*** |
| | | | | (0.0175) |
| Pat/RDC | | | | 0.00348** |
| | | | | (0.00144) |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 87,800 | 87,800 | 87,799 | 87,799 |
| avg. R-squared | 0.186 | 0.185 | 0.178 | 0.178 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

**Table J2: Tobin's Q as a function of disaggregated Innovation productivity and efficiency variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.0379 | 0.0380 | 0.0366 | 0.0365 |
| | (0.0249) | (0.0249) | (0.0253) | (0.0254) |
| RDBE | 0.0917*** | 0.0901*** | 0.0878*** | 0.0877*** |
| | (0.0180) | (0.0184) | (0.0189) | (0.0189) |
| Pat_clean/Book | 1.919* | | | |
| | (0.947) | | | |
| Pat_dirty/Book | -0.511 | | | |
| | (1.346) | | | |
| Pat_other/Book | 0.184*** | | | |
| | (0.0607) | | | |
| Pat_grey/Book | 2.105 | | | |
| | (1.433) | | | |
| Pat_emtech/Book | 0.362** | | | |
| | (0.150) | | | |
| Cit/Book | 0.0103 | | | |
| | (0.00862) | | | |
| Cit_clean/Book | | 0.344** | | |
| | | (0.133) | | |
| Cit_dirty/Book | | -0.139 | | |
| | | (0.213) | | |
| Cit_other/Book | | 0.00901 | | |
| | | (0.0103) | | |
| Cit_grey/Book | | 0.329* | | |
| | | (0.179) | | |
| Cit_emtech/Book | | 0.0434** | | |
| | | (0.0192) | | |
| Pat/Book | | 0.158** | | |
| | | (0.0625) | | |
| Pat_clean/RDC | | | 0.0846* | |
| | | | (0.0449) | |
| Pat_dirty/RDC | | | -0.00890 | |
| | | | (0.0504) | |
| Pat_other/RDC | | | 0.000703 | |
| | | | (0.00110) | |
| Pat_grey/RDC | | | 0.326* | |
| | | | (0.166) | |
| Pat_emtech/RDC | | | 0.0629*** | |
| | | | (0.0163) | |
| Cit/RD | | | 0.00142*** | |
| | | | (0.000363) | |
| Cit_clean/RD | | | | 0.0286** |
| | | | | (0.0109) |
| Cit_dirty/RD | | | | 0.000913 |
| | | | | (0.0112) |
| Cit_other/RD | | | | 0.00194*** |
| | | | | (0.000658) |
| Cit_grey/RD | | | | 0.0541*** |
| | | | | (0.0173) |
| Cit_emtech/RD | | | | 0.0136*** |
| | | | | (0.00440) |
| Pat/RDC | | | | 0.00330** |
| | | | | (0.00145) |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 87,800 | 87,800 | 87,799 | 87,799 |
| avg. R-squared | 0.187 | 0.186 | 0.180 | 0.179 |

Notes. The Table presents the regression results of various specifications (columns 1 and 2) of the Model $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/Book_{it} + \gamma_3 Cit^*/Book_{it} + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ and the Model (columns 3 and 4) $logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Pat^*/RDC_{it} + \gamma_3 Cit^*/RD_{it} + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$ that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to clean patents and clean citations and tests the efficiency in the knowledge creation process, from investment in R&D to efficiency of R&D investment in generating clean patents and citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity; Pat/RDC as a proxy for patent efficiency; and Cit/RD as a proxy for citation efficiency. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. All the variables are defined in Table 1 and we use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

### 2.8.11 Internet Appendix K: US listed firms

**Table K1: Tobin's Q as a function of disaggregated Innovation citation productivity variables, estimated using non-linear least squares method**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.100 | 0.0997 | 0.0997 | 0.0997 |
| | (0.212) | (0.212) | (0.212) | (0.212) |
| RDBE | 0.334*** | 0.203** | 0.243** | 0.200* |
| | (0.0958) | (0.0772) | (0.0869) | (0.0789) |
| Cit_clean/Book | 0.340* | 0.298* | 0.281* | 0.296* |
| | (0.135) | (0.122) | (0.120) | (0.122) |
| Cit_dirty/Book | -0.716** | -0.737*** | -0.762** | -0.745*** |
| | (0.251) | (0.217) | (0.249) | (0.225) |
| Cit_other/Book | | 0.0827** | | 0.0798* |
| | | (0.0320) | | (0.0373) |
| Pat/Book | | | 0.334* | 0.0228 |
| | | | (0.146) | (0.138) |
| Time FE | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 18556 | 18556 | 18556 | 18556 |
| Adjusted $R^2$ | 0.161 | 0.164 | 0.162 | 0.164 |

Notes. The Table presents the regression results of various specifications of the Model $logQ_{it} = \alpha + log(1 + \gamma_1 RDBE_{it} + \gamma_2 Cit\_clean/Book_{it} + \gamma_3 Cit\_dirty/Book_{it} + \gamma_4 Cit\_other/Book_{it} + \gamma_5 Pat/Book_{it} + \sum_{i=1996}^{2012} \kappa_i year_i + \sum_{j=2}^{48} \beta_j Industry_j) + \varepsilon_{it}$ that are estimated using non-linear least squares method and are in the vein of the Models reported in Hall et al., 2005. These Models test whether the knowledge creation process acts as a continuum from R&D to patents and clean citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity. Our dependent variable is the natural logarithm of Tobin's Q and we report clustered standard errors in parentheses. We use the following significance stars $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

**Table K2: Tobin's Q as a function of disaggregated Innovation citation productivity variables, estimated using Fama-MacBeth regressions**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Intercept | 0.450*** | 0.496*** | 0.601*** | 0.570*** |
| | (0.135) | (0.125) | (0.129) | (0.130) |
| RDBE | 0.136*** | 0.0988*** | 0.108*** | 0.106*** |
| | (0.0288) | (0.0237) | (0.0272) | (0.0240) |
| Cit_clean/Book | 0.199*** | 0.180*** | 0.169** | 0.170** |
| | (0.0574) | (0.0563) | (0.0630) | (0.0720) |
| Cit_dirty/Book | -0.774* | -0.805* | -0.800* | -0.857* |
| | (0.411) | (0.407) | (0.413) | (0.426) |
| Cit_other/Book | | 0.0230*** | | 0.0212 |
| | | (0.00635) | | (0.0131) |
| Pat/Book | | | 0.106*** | 0.0323 |
| | | | (0.0237) | (0.0640) |
| Industry FE | YES | YES | YES | YES |
| Firm-level controls | NO | NO | NO | NO |
| Observations | 18,556 | 18,556 | 18,556 | 18,556 |
| avg. R-squared | 0.174 | 0.178 | 0.175 | 0.180 |

Notes. The Table presents the regression results of various specifications of the Model

$$logQ_{it} = \alpha + \gamma_1 RDBE_{it} + \gamma_2 Cit\_clean/Book_{it} + \gamma_3 Cit\_dirty/Book_{it} + \gamma_4 Cit\_other/Book_{it} + \gamma_5 Pat/Book_{it} + \sum_{j=2}^{48} \beta_j Industry_j + \varepsilon_{it}$$

that are estimated using Fama-MacBeth method. These Models test whether the knowledge creation process acts as a continuum from R&D to patents and clean citations. In our specifications we use RDBE as a proxy for R&D productivity; Pat/Book as a proxy for patent productivity; Cit/Book as a proxy for citation productivity. Our dependent variable is the natural logarithm of Tobin's Q and we report standard errors in parentheses. We use the following significance stars *** p<0.01, ** p<0.05, * p<0.1.

# National culture 'profiling' in machine-learning applications: The utility and ethics of applying value ascriptions in global alert models

**Abstract**

We examine the utility of incorporating national culture profiling in bank-level machine-learning informed alert models, which relate to financial malfeasance. At a globally important financial institution, we use binary classifier type alert models and establish the utility of dimensions of national culture in formulating anti-money laundering predictions. For corporate (individual) accounts, Hofstede individuality (individuality, and national-level corruption perception and financial secrecy) scores of the country in which a customer is resident, or from which a wire is sent/received, are of paramount importance. When combined with extensive account and transaction data; as well as even a proprietary institutional algorithm, national culture traits markedly enhance the models' predictive performances. We consider the ethical implications of ascribing values, against a global standard, to dimensions of national culture. We offer an ethical framework for the use of national profiling in anti-fraud alert models.

**JEL Classification: C52, C55, D12, G17, G21**
**Keywords:** National Culture Profiling, Machine Learning, Anti-Money Laundering

## 3.1 Introduction

Pervasive across borders and undermining local economies, money-laundering remains an issue of global concern. A channel to legitimize dirty money (i.e., money generated from illegal activities), it integrates such monies into an established financial system for subsequent use without evoking suspicion (FATF, 1999; IMF, 2021).[42] In facilitating the generation and disbursement of illicit proceeds from criminal activities, it paves the way for further financial illegal activity, compounding the problem. The upshot of money-laundering is hence the perpetuation of associated crime, the misallocation of capital and the possibility of international financial instability. Although difficult to measure, estimates for the total amount of money-laundered worldwide range from 2-5% of global GDP (approximately $600 billion to $1.6 trillion) (UN, 2020). Since the financial sector is critical to the transmission of shocks in the real economy, financial misconducts in the banking sector can have widespread repercussions (Cornett et al., 2011; Vinas, 2021).[43]

Current anti-money laundering (AML) surveillance is painstakingly inefficient, time-consuming,

[42]Please see Internet Appendix C for an overview on money-laundering, the efforts of financial institutions to combat this fraudulent practice, and the existing inefficiencies in AML surveillance.

[43]For instance, financial misconducts can undermine trust in financial institutions and markets, create systemic risks, besides potentially harming consumers (Financial Stability Board, 2018).

and labor-intensive. Financial institutions vet thousands of potentially suspicious transactions every day and any failure to comply with the AML surveillance requirements often makes them liable for substantial fines and penalties levied by the regulatory bodies. Since the volume of banking and transactional data have increased exponentially in recent times, financial institutions are increasingly applying machine learning to detect and curb money-laundering (FATF, 2021). While these machine learning alert models hinge on proprietorial data (FATF and Egmont Group, 2020), we seek to address, in our paper, whether publicly available data on national culture, in particular, has the predictive capacity to detect money-laundering.

In this paper, we investigate the relevance of several country-specific culture and institution quality indices, relative to the account- and transaction-level information on the financial institution's customers, against modelling the incidence of suspicious money movement within a financial institution. In other words, we examine whether national culture impacts a bank customer's predilection for bank fraud. We ask, in particular, whether individualism scores (Hofstede, 2001) pertaining to a customer's country of residence and/or the country of wire origination/destination are useful in detecting money-laundering in our models accounting for the financial institution's proprietary data.

Our research question in addressing recent literature examines the utility of incorporating national culture profiling in bank-level machine learning informed alert models for detecting money-laundering at a globally prominent financial institution. However, our research question departs from the existing literature in examining if banking customers' socio-cultural matrix inspires their predilections for committing money-laundering instead of examining the bank's corporate culture or its employees' nationality. Our study provides insights and empirical evidence for financial institutions willing to benefit from incorporating machine learning and publicly available data to their existing data framework to enhance AML operation.

In light of recent literature on the role of culture in corporate misconduct and bank failure (Berger et al., 2019; Liu, 2016; DeBacker et al., 2015; Bame-Aldred et al., 2013), we explore the relevance of several country-specific culture and institution quality indices vis-à-vis modelling incidence of suspicious money movement within a financial institution. As individuals may not always hold unbiased beliefs and can behave irrationally (Kim et al., 2016), the anticipated incentives and deterrents for misconduct and the anticipated likelihood of being held accountable for wrongdoing, can vary substantially across national cultures (Husted, 2000). The social normativity of national culture (Goodell, 2019), in particular, can influence misconduct among the customers of financial institutions.[44] We further assess the importance of national culture traits relative to customers' account and transaction traits. In so doing, this paper investigates if a banking customers' socio-cultural matrix inspires their predilections for

[44]National culture also figures prominently in assessing ethical values and discernment in business ethics research (Armstrong, 1996; Davis and Ruhe, 2003; Getz and Volkema, 2001; Vitell et al., 1993; Volkema, 2004).

committing money-laundering.

To investigate our research questions, we employ a major global financial institution's large proprietary dataset containing cross-border wire transactions made during 2009-2018. The dataset pertains to alerts generated by international wire transfers both to and from customers of the institution. The financial institution's monitoring system generates an alert for a wire transfer, if the wire amount exceeds a predetermined threshold and if the country from which the wire is sent and/or received falls in the list of countries blacklisted by the financial institution. The alert is then investigated by a team of experts. In their judgement, if the corresponding wire transaction seems highly suspicious, then they escalate it to an issue case and refer the matter to higher authorities for further investigation. The issue cases can be regarded as precursors to money-laundering. We further collate the novel proprietorial customer- and account-level cross-border wire transfer bank client data with country-specific culture (Hofstede's cultural dimensions) and institution quality indices (Corruption Perception Index; Financial Secrecy Index).[45] Since the proprietorial dataset provides a clearly labelled response variable (Issue Case), we employ supervised learning techniques such as logistic regressions, random forest, gradient boosted machines, and support vector machines to detect money-laundering at the financial institution.

Individualism is linked to behavioral attributes of over-confidence and self-attribution bias (Chui et al., 2010; Heine, 2003; Li et al., 2013; Markus and Kitayama, 1991; Pfeffer and Fong, 2005). Due to these behavioural attributes, the individuals concerned show low levels of self-monitoring (Biais et al., 2005), and are over-optimistic in respect to the precision of their predictions (Van den Steen, 2004). We, therefore, expect that bank customers, in more individualistic countries, can overestimate their abilities (Heine et al., 1999; Markus and Kitayama, 1991) to opportunistically (Chen et al., 2002) disguise misconduct so that financial institutions will not detect their behavior. We, therefore, examine if a banking customers' predilections for committing money-laundering can be due to cross-country cultural differences linked to that facet of national culture known as individualism.

Our empirical work is supported by a dual process understanding of an individual's cognition, personal attitudes, and values (Fischer et al., 2010; Peterson and Barreto, 2018), on the one hand, and, apprehending the societal culture facets that also inform the cognition of individuals regarding opportunities for financial misconduct (Peterson and Barreto, 2018; Watts et al., 2020), on the other. In line with recommendations in Kirkman et al. (2006), we examine whether individualism as a trait (Hofstede, 2001) is useful in informing our bank-level machine learning informed alert models relating to financial malfeasance that accounts for pertinent country-, account- and transaction-level features of the financial institution's clients. We

[45]We employ Hofstede (2001) individualism, masculinity, power-distance, and uncertainty avoidance, national culture dimensions, inspired by prior literature. We also employ two institution quality indices, namely, corruption perception index and financial secrecy index to measure the levels of corruption and financial secrecy of a country.

provide a brief outline of our findings.

We find country-level factors, particularly national culture as comprising strong predictors of identifying suspect bank wire transfers. Using binary classifier type alert models, together with corrections for data imbalance, our results reflect the strength of national culture dimensions in formulating anti-money laundering predictions. For corporate accounts, Hofstede individuality scores of the country in which a customer is resident, and from which a wire is sent/received are the most important factors. For individual accounts, individuality scores of the country in which a customer is resident; national-level corruption perception scores of the country in which a customer is resident, and from which a wire is sent/received; and financial secrecy scores of the country in which a customer is resident are the most important factors. National culture alone provides a high degree of predictive power. And when combined with extensive account and transaction data, its inclusion greatly enhances predictive ability. For instance, for corporate-related alerts, individuality rating of the customer's country of residence, individuality rating of the country of wire origination/destination, and the uncertainty avoidance cultural trait of the customer's residence country rank among the top five features to assess the customer's predilections for money-laundering. For people-related alerts, the individuality score of the customer's residence country, corruption perception score of the country of wire origination/destination, and financial secrecy score of the customer's country of residence are the most important county-level features that rank among the top ten features. We further enlarge our feature space to include proprietary risk score assigned to the alerts by the financial institution's proprietary algorithm to examine the relative importance of country-level features in detecting money-laundering. We find that for corporate-related alerts, individuality rating of the country of wire origination/destination, individuality rating of the customer's country of residence, and corruption perception score of the country of wire origination/destination have higher predictive capacity than the proprietorial risk score. However, in case of people-related alerts, the proprietorial risk score is the most important feature. Overall, our results suggest that the inclusion of country-level features greatly enhance the predictive ability of our models in detecting money-laundering. Pertinently, our findings provide practical implications for the financial services sector in terms of AML compliance and prevention strategy.

More broadly, this paper contributes to the literature that investigates the determinants of financial malfeasance. Financial misconducts can undermine trust in financial institutions and markets, create systemic risks, besides potentially harming consumers (Financial Stability Board, 2018). Given the magnitude of the problem, several studies have investigated the determinants of financial malfeasance. For instance, Efendi et al. (2007) find that if a firm's CEO has substantial holdings of in-the-money stock options, then it increases the likelihood of misstated financial statements. Further, the authors observe that firms constrained by an interest-rate debt covenant, firms having CEOs serving as the board chair, and firms raising new debt or equity capital are more likely to misstate their financial statements. Dimmock and Gerken (2012) em-

ploy a panel of mandatory disclosures filed with the SEC to test the predictability of investment fraud. They observe that conflicts of interest, disclosures related to past regulatory and legal violations, and monitoring significantly predict fraud. Interestingly, recent literature investigates whether the cultural background of banking employees and their personal attributes influence their predilection for opportunistic behavior. For example, Liu (2016) investigates whether a CEO's ancestry affects the likelihood of financial misconduct. Parsons et al. (2018) examine whether geography-based social norms could impact misconduct. Davidson et al. (2015) and Griffin et al. (2017) probe whether unethical behavior in their personal lives is likely to lead to financial misreporting in their role as managers of firms.

The papers that are closely related to our work include those of Liu (2016), DeBacker et al. (2015) and Fisman and Miguel (2007). All these papers reveal that culture can influence an individual's decision for wrongdoing. Liu (2016) investigates whether the cultural background of key employees of a firm in influencing their opportunistic behavior impacts corporate misconduct. He constructs a measure of corporate culture, corporate corruption culture, using data on the cultural background of officers and directors of a firm to proxy for a firm's opportunistic behavior. He detects that firms with high corruption culture are more likely to engage in corporate misconduct. His findings show that when individuals emigrate their propensity for financial wrongdoing is influenced by the culture of their country of ancestry. DeBacker et al. (2015) examine the impact of culture on corporate behavior. They show that corporations in the US headed by foreign nationals belonging to countries with higher corruption norms evade more tax. They also find that enforcement measures to increase tax compliance were less effective in deterring tax evasion of these corporations. Fisman and Miguel (2007) found that United Nations' diplomats from countries prone to high-corruption exhibited a greater propensity to engage in illegal parking in New York City prior to 2002 when diplomatic immunity protected the UN diplomats from parking enforcement actions. This diplomatic immunity norm provided the authors grounds to conduct a natural experiment for testing the role of cultural norms on their illegal parking behavior.

Our findings provide practical implications for the financial services sector in terms of AML compliance and prevention strategy. Confirming the conduciveness of machine learning in incorporating national culture, the findings also contribute to the extensive literature that ascribes values to ethicality and discernment constituting distinct national traits. The use of demographic inputs, particularly country-level factors in machine learning models, touches on a wide variety of literature that incorporate cultural and demographic variables that ascribe value to these characteristics.[46] We offer, a framework for assessing the ethics of using country-level factors in machine learning prediction and detection. We identify several characteristics of the use of country-level factors in machine learning procedures that are central to evaluating the

[46]See, for instance, the area of microfinance, where the notion of female borrowers being more trustworthy undergirds the industry (Aggarwal et al., 2015)

ethics of their respective usage. These include: 1). Do public good concerns in countering money-laundering outweigh 'collective treatment' in algorithmic national profiling? 2). Do those issuing the alerts permitted to avail of the personal data? 3). Who is responsible for the design of an algorithm? 4). Are algorithms accountable? 5). Are the algorithms used for detection or prediction? And are there subtle distinctions between them? 6). Do alert models reflect global, national, or sub-national, public, or private regulation? 7). Do the existing algorithms exacerbate tangential social biases?

The rest of this paper is organized as follows. Section 2 presents literature review. Section 3 presents our testable hypotheses. Section 4 discusses the proprietary dataset, country-specific culture, and institution quality indices from which we have drawn our predictors/features. Section 5 outlines the various data resampling methods used in the paper for meaningfully sourcing information from the data, and discusses the machine learning methodologies, performance evaluation metrics, and feature importance metrics of the study. Section 6 presents the empirical findings. Section 7 provides a framework for evaluating the ethics of machine learning prediction and alert models. Finally, section 8 concludes.

## 3.2 Literature Review

We examine whether national culture traits profiling can usefully inform a machine learning alert model in detecting money-laundering at a globally prominent financial institution. In light of recent literature on the role of culture in corporate misconduct and bank failure (Berger et al., 2019; Liu, 2016; DeBacker et al., 2015; Bame-Aldred et al., 2013), we explore the relevance of several country-specific cultural and institution quality indices vis-à-vis modelling incidence of suspicious money movement within a financial institution. Our rationale for employing national culture as a predictor of bank fraud is further borne out by prior literature that correlates national culture dimensions with the quality of ethical behaviour and perception (Armstrong, 1996; Davis and Ruhe, 2003; Getz and Volkema, 2001; Vitell et al., 1993; Volkema, 2004). For instance, Vitell et al. (1993) posit a conceptual framework that employs Hofstede's cultural dimensions to understand the influence of culture on ethical decision-making. As culture represents a society's shared beliefs, values, and ideals, it follows why national culture dimensions are able to explain ethical decision-making. Therefore, our study investigates if a banking customers' socio-cultural matrix informs their predilections for committing financial misconduct, namely, money-laundering. To answer this question, we employ Hofstede (2001) individualism, masculinity, power-distance, and uncertainty avoidance, national culture dimensions, inspired by prior literature. We also employ two institution quality indices, namely, corruption perception index and financial secrecy index to measure the levels of corruption and financial secrecy of a country.

### 3.2.1 Association of national culture with quality of ethical behaviour and perception

Prior studies associate national culture dimensions with the quality of ethical conduct and perception. For instance, Vitell et al. (1993) propose a theoretical framework to understand the association between culture and ethical decision-making in business. They argue that in individualistic societies, business practitioners are less likely to adhere to both formal codes of ethics and informal norms than their counterparts in countries that value collectivist ethos. This stands to reason since individualistic societies that value self-reliance, freedom, and achievement, tend to consider its people's actions as beyond reproach. The authors also underscore that business practitioners in a more hierarchical society (high score on Hofstede's power-distance index) tend to draw ethical cues from their superiors rather than peers, and that they value formal ethical codes than informal norms to fashion their own deontological norms. In contrast, business practitioners in societies with low power-distance are more likely to draw ethical cues from peers and consider informal norms more important than formal ethical codes in forming their own deontological norms. The authors further maintain that in masculine cultures, business practitioners are less likely to perceive an ethical violation than their counterparts in feminine cultures because their cultures do not define such violations in relation to ethics. Besides, in masculine cultures, business practitioners are less likely to be influenced by formal ethical codes than their counterparts in feminine cultures. Vitell et al. (1993) postulate that business practitioners from societies scoring high on uncertainty avoidance are more likely to consider formal ethical codes to form their deontological norms as well as be mindful of negative consequences of questionable actions than their counterparts. More pertinent, business practitioners in high uncertainty avoidance countries are less likely to perceive ethical concerns than their counterparts.

Likewise, Volkema (2004) discovers that people in individualistic, masculine, and high uncertainty avoidance societies are much more likely to adopt competitive and questionable negotiation behaviors than people in collectivist societies. Contrarily, people in high power-distance countries are less likely to opt for competitive and questionable negotiation practices.

Getz and Volkema (2001) report that in high power-distance cultures, both high-level public officials and members of the underclass are more prone to unethical behavior (bribery, extortion). In highly hierarchical and unequal societies, high-ranking officials often exploit their class privileges for personal gains; conversely, members of the underclass take recourse to unethical practices because it helps them to improve their standard of living. The authors also note that bribery and corruption are more likely to occur in masculine rather than feminine cultures. This is because masculine cultures that subscribe to the ends justifying the means approach measure achievement in terms of material success.

Aggarwal et al. (2014) examine the association between national culture dimensions and cross-national differences in Graduate Management Admission Test (GMAT) scores. They find a

negative association between mean GMAT scores with masculine and high power-distance culture dimensions. Further, the authors observe a positive association between a country's mean GMAT score with its level of uncertainty avoidance and individualistic traits. They conclude, on an average, 80-point difference in cross-national mean GMAT scores could be explained by the cultural factors.

### 3.2.2 Association of national culture with finance-related behaviour

Literature associates national culture with a host of finance-related behavior. For instance, Chui et al. (2010) investigate whether cross-country cultural differences influence returns of momentum strategies. They find individualism positively associated with trading volume, volatility, and magnitude of momentum profits. Lievenbrück and Schmid (2014) examine if cross-country cultural differences could explain firms' hedging decisions. They report that a country's long-term orientation significantly lowers its propensity to hedge; likewise, firms in countries that exhibit a high masculine trait are less likely to hedge using options. Aggarwal and Goodell (2009) in their study of national preferences for financial intermediation show that national culture strongly influences a country's preference for financial intermediation (markets versus institutions). Chui et al. (2002) claim that national culture impacts choice of firm leverage. Shao et al. (2013) investigate whether individualism affects firms' ambition to undertake investments. They find that firms in individualistic countries invest more in long-term risky assets and R&D projects than short-term safe and physical assets. Further, they note that firms in individualistic countries employ excess cash to invest in R&D rather than increasing their dividends. Further, Aggarwal and Goodell (2013) explore whether national culture could explain the differences in pension systems, internationally. They detect a negative association between pension progressivity with masculinity, uncertainty avoidance, individualism, long-term orientation, employment rights, average pension levels, social trust and economic inequality.

## 3.3 Hypothesis Development

We use national culture facets (i.e., societal cultural value dimensions - as in Peterson and Barreto (2018)), in the first instance, to indicate the societal context of individuals.[47] In so doing, we do not require assignment of societal characteristics to individuals to replace omitted information about their personal values (Kirkman et al., 2017; Tung and Stahl, 2018). Instead, we highlight that the contextual effects of societal institutions and norms can inform "individuals experience and, hence, what people unconsciously intuit and consciously understand" (Peterson and Barreto, 2018; Goodell, 2019). This can, in turn, contribute to an individual's

[47]In line with argumentation in Peterson and Barreto (2018), national culture facets can indicate contextual characteristics that more strongly shape an individual's cognition, than do consciously expressed personal values. For instance, it can be comparatively insightful to know of the societal context in which an individual was raised, as opposed to her self-professed attitudes and values, with a view to predicting, in a corrupt system, her interaction with regulations.

cognition and decisions. In this way, we use national culture facets to represent the context in which a society's members react to culture. This study, more specifically, considers if a banking customer's opportunity and inclination to commit financial misconduct – laundering money, is informed by her cultural context.

We note, however, in the second instance, that national culture facets can approximate values at the individual level. While the assignment of country level scores to our samples necessarily neglects within-country variability (Kirkman et al., 2006), Fischer et al. (2010) report strong evidence of the structural similarity of values at the individual and country levels. Such values data can serve as a robust approximation for the cultural knowledge, resources, structures and norms held by society (Peterson and Barreto, 2018). Using this line of rationale, Westjohn et al. (2021) show that cultural values moderate the relationship between consumer animosity, against a foreign brand (for whatever reason), and an individual's willingness to buy. In the same vein, Watts et al. (2020) show that high levels of uncertainty avoidance strengthen the relationship between transformational leadership and employee innovation.

As a result, we state two channels by which national culture facets can influence the decision making of an individual.[48] This representation can be viewed as a dual process understanding of cognition for an individual: personal attitudes and values, at one level, together with societal culture facets, at another level, informing an individual's cognition.

Due to one or both of these channels, we, therefore, infer that banking service clients' propensity toward malfeasance can vary markedly across national cultures. As individuals may not always hold unbiased beliefs and can act irrationally (Kim et al., 2016), the anticipated incentives and deterrents for misconduct and the anticipated likelihood of being held accountable for wrongdoing, can vary substantially across national cultures (Husted, 2000). The social normativity of national culture (Goodell, 2019), in particular, can influence misconduct among the customers of financial institutions.

### 3.3.1 Individualism

Hofstede's individualism index provides insights into people's level of interdependency in a society. In an individualistic society, people are expected to fend for themselves and their immediate families; whereas in a collectivist society, in-groups to which people belong look after them in exchange for unquestioning loyalty. Individualism is linked to behavioral attributes of over confidence and self-attribution bias i.e., people's tendency to attribute positive events to their own character but attribute negative events to external factors (Chui et al., 2010; Heine, 2003; Li et al., 2013; Markus and Kitayama, 1991; Pfeffer and Fong, 2005). Due to these behavioural attributes, such individuals show low levels of self-monitoring (Biais et al., 2005),

[48]Moreover, national culture constructs can be primed and made temporarily accessible (Leung and Morris, 2015), indicating that they do manifest at the individual level.

and they are over-optimistic in respect to the precision of their predictions (Van den Steen, 2004). This can give individuals over-optimistic views of the future (Fischer and Chalmers, 2008) and lead to their inaccurate evaluation of bad news (Kim et al., 2016). As a result, we expect that bank customers, in more individualistic countries, can overestimate their abilities (Heine et al., 1999; Markus and Kitayama, 1991) to opportunistically (Chen et al., 2002) disguise misconduct so that financial institutions and regulators will not detect their behavior.

Further, because individualistic cultures value self-reliance, freedom, achievement, and tend to consider its people's actions as beyond reproach, prior studies argue that individualistic cultures promote ethically questionable behavior (Bame-Aldred et al., 2013; Cullen et al., 2004; Martin et al., 2007; Vitell et al., 1993). Additionally, individualism is also linked to risk-taking behavior (Gaganis et al., 2019; Mourouzidou-Damtsa et al., 2019).

Also, Chui et al. (2010) and Kreiser et al. (2010) suggest that in more individualistic countries, decisions, in general, are more likely to be taken by individuals rather than the group. In such countries, people have a strong belief in individual choices and decisions (Markus and Kitayama, 1991). This is critical as, according to Shupp and Williams (2008), in high risk situations, individuals are more risk tolerant than groups. Similarly, the incentive to perform in individualistic countries is underpinned by compensation practices that focus on individual recognition (Schuler and Rogovsky, 1998). As incentivized and risky decision making is more likely in individualistic cultures, we expect that banking misconduct can also be more likely in this setting.

Collectively, the above arguments suggest that a banking customers' predilections for committing money-laundering can be due to cross-country cultural differences linked to that facet of national culture known as individualism. In other words, individualism scores pertaining to a customer's country of residence and/or the country of wire origination/destination are useful in detecting money-laundering, when our models employ only the country-level variables. Our initial major hypothesis can, thus, be stated:

**Hypothesis 1**: Individualism is useful in detecting money-laundering when our models employ only the country-level features.

However, prior literature also identifies a countervailing outcome between individualistic cultures and its peoples' ethically questionable behavior. For instance, some studies find a negative association between individualism and tax evasion (Bame-Aldred et al., 2013; Tsakumis et al., 2007; Richardson, 2008). Further, Cullen et al. (2004) detect a negative relationship between individualism and managers' ethically questionable decisions. Similarly, Armstrong (1996) notes a positive association between individualism and higher ethical standards. Therefore, one can put forward an alternative hypothesis that the individualism score of a bank customer's country is negatively related to her predilection for money-laundering. It is, therefore, a moot

question whether individualism is associated with fraud.

We, further, examine whether individualism is useful in informing our models' outcomes when we account for the account- and transaction-level features of the financial institution's clients. Existing evidence shows that financial institutions use customers' account- and transaction-level information to monitor suspicious money transfers (FATF and Egmont Group, 2020). As a result, we extend our dataset to include account- and transaction-level features of the financial institution's clients, and examine whether the individualism traits pertaining to a customer's country of residence and/or the country of wire origination/destination are useful in detecting money-laundering.

**Hypothesis 2**: Individualism is useful in detecting money-laundering when our models employ the country, account, and transaction level features.

We also seek to investigate the relative importance of individualism when we add the proprietary risk score, PROP Score, to our models containing country-, account-, and transaction-level features.[49] It is reasonable to surmise that the financial institution's proprietary algorithm employs data to which we do not have access to. To examine the predictive capacity of individualism trait relative to the financial institution's proprietary alert algorithm's risk scores, we include PROP Score to our dataset.

**Hypothesis 3**: Individualism is useful in detecting money-laundering when our models employ the country-, account-, transaction-level features along with the proprietorial risk score.

## 3.4 Data

This study employs a major global financial institution's large proprietary dataset consisting of cross-border wire transactions made during 1 January 2009- 31 December 2018. The data pertains to alerts generated by international wire transfers both to and from customers of that institution. An alert is generated for a wire transfer in the financial institution's monitoring system, if the wire amount exceeds a predetermined threshold and if the country from which the wire is sent and/or received falls in the list of countries blacklisted by the financial institution. The alert is then investigated by a team of experts. In their judgement, if the corresponding wire transaction seems highly suspicious, then they escalate it to an issue case and refer the matter to higher authorities for further investigation. Alerts are generated for more than 60,000 customer accounts in the data provided by the financial institution. These accounts can be broadly classified into six account registration types.[50] Among the six account registration types, we focus only on two, the corporate- and people-related account registrations, since these pertain to 78.23% of the alerts and 93.77% of the issue cases. Table 1, Panel A reports the number of

[49] PROP Score is the risk score assigned to the alerts by the financial institution's proprietary alert algorithm.
[50] See Internet Appendix A.

alerts and subsequent issue cases over time, associated with the corporate- and people-related accounts. The total number of alerts generated during the ten-year period is 206,751. In considering only the corporate- and people-related account registration types, however, the total number of alerts reduces to 153,917.[51]

### 3.4.1 Sample Selection

The dataset provides information on the wire transactions that generated the alerts and the corresponding customers' accounts and transaction history. However, we do not have complete information on customers' account- and transaction-level data. Considering this deficiency, only 60% of the alerts could be matched to the wire transactions that triggered the alerts. Table 1, Panel B reports the number of alerts and subsequent issue cases that can be successfully matched with the corporate- and people-related account registration types.

[Please insert Table 1 about here.]

### 3.4.2 Dependent Variable

The dependent variable in this study is the outcome of an investigation – specifically, whether an alert is deemed to be highly suspicious, i.e., an issue case. Alerts are generated through an automated process based on a customer's aggregated wire transactions exceeding a certain threshold and whether the wires involved any blacklisted countries, on a given day. The alerts are then examined by a team of investigators. Each alert passes through several phases of escalation before reaching the status of issue case. It is only then that the case is passed on to higher authorities for legal processing.

Far from efficient, this method of screening transactions for suspicious activity remains standard across the industry, since financial governing bodies enforce harsh penalties on institutions that they deem to be lax in detecting money laundering.[52]

### 3.4.3 Feature Selection

In creating features from the data of the financial institution, we account for customers' account and transaction history. However, our chief novelty consists in creating features from country-specific culture and institution quality indices which go into investigating if banking customers' socio-cultural matrix influences their predilections for committing financial misconduct, namely, money-laundering. We further group the features into three categories: (1) Country-level, (2) Account-level, and (3) Transaction level. Below, we discuss the features included in our study. Concise definitions are provided in Table 2.

[51] See Panel B in Table 1.

[52] The proportion of false alarms typically exceeds 99%. To avoid confusion, we reserve the use of the term "false positives," for reference to the model results.

[Please insert Table 2 about here.]

#### 3.4.3.1 Country-level Predictors

We create quantifiable country-level features from internationally recognised country-specific culture and institution quality indices. Corresponding to each of the indices, we create two features which we distinguish as the origin/destination country of the wire transaction and the residence country of the customer receiving/sending the wire that triggered an alert. We create two sets of features for each index, since in the financial institution's data a customer's residence country is documented precisely, though often the data is unclear on the country to/from which the customer is sending/receiving the wire.[53] Thus, we distinguish the features constructed from a particular index by using the subscripts R and W; where R and W denote customer's country of residence and the country of wire origin/destination, respectively. Below, we define country-specific culture indices and institution quality indices employed to create features in our study. In this study we employ four national culture dimensions proposed by Hofstede (2001) and two internationally recognized indices that measure the levels of corruption and financial secrecy of a country.

1. **Individualism Index** $(IDV_R, IDV_W)$
   This Hofstede index provides insights into the level of people's interdependency in a society. In an individualistic society, people are expected to fend for themselves and their immediate families; whereas in a collectivist society, in-groups to which people belong look after them in exchange for unquestioning loyalty. A country scoring high on this index exhibits individualistic trait, whereas a country with a low score cherishes a collectivist ethos.

2. **Masculinity Index** $(MAS_R, MAS_W)$
   This Hofstede's culture dimension quantifies the extent to which a society values achievement, success, and competition (masculine traits) over modesty and compassion towards others (feminine traits). A country scoring high on this index indicates that its people privilege masculine over feminine traits.

3. **Power-Distance Index** $(PDI_R, PDI_W)$
   The index provides insights on the level of inequality endorsed and accepted by the less powerful members of a society. A country with a low score on this index shows its citizens as having a lower tolerance for social inequality and vice-versa.

4. **Uncertainty Avoidance Index** $(UAI_R, UAI_W)$
   A Hofstede's culture dimension, the index quantifies the impact of national culture on its

[53] For instance, in some wire transfers the IBAN of the customer sending/receiving the wire is documented, whereas in others we retrieve this information from the address/information field documented in the data.

peoples' tolerance to deal with uncertainty. Cultures that try to minimize ambiguity rank high on this index and vice-versa.

5. **Corruption Perception Index** $(CPI_R, CPI_W)$
   Drawing on thirteen different data sources, Transparency International's composite index ranks countries/territories based on the perceived corruption in public sector by experts in governance and business climate analysis. The index ranks 180 countries on a scale of 0 to 100, where 0 corresponds to high perceived level of corruption and 100 to low perceived level of corruption, respectively.[54]

6. **Financial Secrecy Index** $(FSI_R, FSI_W)$
   Proposed by Tax Justice Network, the index ranks jurisdictions based on the scale of their offshore financial activities and the regulatory framework providing legal and financial secrecy to businesses and individuals based elsewhere. The index provides insights on global financial secrecy, tax havens, and illicit financial flows (Puspitasari, Sukmadilaga, Suciati, Bahar, and Ghani, Puspitasari et al.; Houqe et al., 2015; Michalos and Hatch, 2019; Hassan and Giorgioni, 2015).[55]

#### 3.4.3.2 Account-level Predictors

We employ four account-level features from the financial institution's dataset, namely *Customer Age*, *Account Age*, *Customer Net Worth*, and *Alert Supplier Code*. The feature *Customer Age* is defined as the age of the customer, a private individual, on the date when the alert is generated. We include this feature when detecting money-laundering exclusively for people-related accounts. Similarly, *Account Age* is the length of time an account stood registered from the date of alert. The feature *Customer Net Worth* is the aggregate balance on all the accounts of the customer. The *Alert Supplier Code* records the type of systemic method that the financial institution employs to collect alerts.

#### 3.4.3.3 Transaction-level Predictors

For each wire transaction that triggered an alert, we have information on, over a 180-day period preceding the alert, the number of incoming and outgoing wires, and transfers to and from the corresponding customer's account $(TFI_{180}, TFO_{180})$; the aggregated amount of incoming and outgoing wires and transfers $(\sum TFI_{180}, \sum TFO_{180})$; the number of incoming and outgoing checks $(CKI_{180}, CKO_{180})$, and the aggregated amount of incoming and outgoing checks $(\sum CKI_{180}, \sum CKO_{180})$.

[54] Liu (2016) indicates that the Transparency Internationals Corruption Perception Index can "capture a general attitude toward opportunistic behavior at the country level."

[55] Hassan and Giorgioni, 2015 indicate that the Tax Justice Network's Financial Secrecy Index at country-level "indicates a lack of transparency and unwillingness to engage in effective information exchange, which makes a secretive country a more attractive location for routing illicit financial flows and for concealing criminal and corrupt activities."

## 3.5 Methodologies

In this section of the paper, we discuss the various data resampling methods for meaningfully inferring information from the data. It then focuses on the machine learning methodologies and the performance evaluation metrics used to evaluate them. Finally, it discusses the feature importance method. We discuss the data resampling techniques in subsection 3.5.1, machine learning methodologies in subsection 3.5.2, performance evaluation metrics in subsection 3.5.3, and feature importance in subsection 3.5.4.

### 3.5.1 Data Balancing

The dependent variable suffers from severe class imbalance. In other words, the number of observations that belong to the positive class (issue case) is significantly lesser than those that belong to the negative class (generated alert is not an issue case). Models trained on such data in prioritizing the prevalent class over the minority class leads to an overly optimistic measure of accuracy (Batista et al., 2004). While such models can detect a non-fraudulent transaction with high level of accuracy, they often fail to detect highly suspicious transactions. Failure to detect highly suspicious transactions could pose a threat to the financial institutions' professional credibility and may also lead to significant regulatory penalties.

In this study, we avail of various data-resampling techniques for overcoming the challenges posed by the imbalanced class distribution. Below, we discuss the resampling techniques employed in our study.

1. **Under sampling:** This technique randomly discards observations from the majority class to better balance the skewed distribution. In reducing the majority class's size to match the minority class, this technique, however, forgoes potentially useful information from the majority class.

2. **Hybrid sampling:** Combining under-sampling and over-sampling methods,[56] this technique applies under-sampling technique to the majority class and over-sampling technique to the minority class to balance the class distribution.

3. **Synthetic sampling:** This technique works like over-sampling. However, instead of randomly duplicating observations from the minority class, it introduces artificial noise to perturb its predictor values to avoid over-fitting. In our study, we use ROSE (Random Over-sampling Examples) synthetic sampling method. This method utilizes the hybrid-sampling technique besides synthetic sampling to overcome the computational challenges of a much larger data set.

[56]Over-sampling: This technique randomly duplicates observations from the minority class to match the majority class size. We refrain from employing this technique as it can be computationally expensive (in cases of severe class imbalance, it may almost double the size of the dataset) and it often leads to overfitting the model.

### 3.5.2 Machine Learning Methodologies

In this subsection, we discuss the machine learning algorithms, namely logistic regression, random forests, support vector machines, and gradient boosted machines, employed in our study to detect money-laundering at the financial institution.

#### 3.5.2.1 Logistic Regression

Logistic regression (LR) models the probability of an observation belonging to a particular class. It employs the logistic function,

$$p(X) = \frac{e^{\beta_0+\beta_1 X_1+...+\beta_p X_p}}{1+e^{\beta_0+\beta_1 X_1+...+\beta_p X_p}} \tag{22}$$

to model the probability of the categorical response variable, Y. In the above logistic function $X_1, X_2, \ldots, X_p$ are the $p$ features. Simple manipulation of the above logistic function gives us,

$$\frac{p(X)}{1-p(X)} = e^{\beta_0+\beta_1 X_1+...+\beta_p X_p} \tag{23}$$

and

$$ln(\frac{p(X)}{1-p(X)}) = \beta_0+\beta_1 X_1+...+\beta_p X_p \tag{24}$$

which shows that the logit, $ln(p(X)/(1-p(X)))$, is a linear function of the features$X_1, X_2, \ldots, X_p$. We estimate the coefficients using the Maximum likelihood method. After the coefficient estimation, we select a suitable probability threshold to classify observations to the two distinct classes. Logistic regression is easy to implement and does not require making assumptions about the class distributions in the feature space. However, since it assumes a linear relationship between the logit and the features, this algorithm fails to more complex non-linear behavior - unless such a relationship is explicitly accounted for.[57]

#### 3.5.2.2 Random Forest

A tree-based machine learning algorithm that in generating multiple decorrelated trees, Random Forest combines their corresponding predictions to arrive at a single prediction. The rationale for this algorithm consists in improving the prediction accuracy vis-à-vis the Decision Tree algorithm. When predictions of several decorrelated decision trees are combined, the resulting machine learning method in registering lower variance leads to better prediction accuracy. Although, the Random Forest achieves higher prediction accuracy than a single decision tree model, it does so at the expense of lower model interpretability. Below, we briefly discuss the decision tree algorithm and then examine the random forest algorithm.

[57] That is, our variables must be transformed accordingly in order capture the behavior we wish to model, e.g., a quadratic or logarithmic function.

Decision Trees

Decision Trees involve stratifying the feature space into non-overlapping regions. For a test observation that falls in a particular region $R_i$, the Decision Tree predicts the response value for the test observation to be the mean or mode (depending on whether the response variable is quantitative or qualitative) of the response values of the training observations in the region $R_i$.

Thus, the recursive binary splitting approach is adopted in constructing the non-overlapping regions of the feature space. This approach popularly known as the top-down greedy approach begins at the top and splits the feature space successively. A feature that results in the highest reduction in the residual sum of squares / classification error rate is considered for a split, at a given step in the tree building process. Each split creates two additional non-overlapping regions. To split one or both the resulting regions, the algorithm chooses the features that minimize residual sum of squares / classification error rate within the regions. This process of splitting ceases when the stopping criterion is met. This approach is called 'greedy' since the feature that minimizes the residual sum of squares / classification error rate the most at a given point privileges a readily available split candidate rather than opting for a feature that could result in a better decision tree in the long-term.

Once the decision tree is developed, for any given test observation, the algorithm first identifies the region to which the test observation belongs. It then assigns the mean/mode of the response values of the training observations belonging to the same region as the response value for the test observation. Although the Decision Tree algorithm is intuitive, unbiased (when grown sufficiently deep), and offers highly interpretable results, being prone to high variance its predictions are often unreliable. The Random Forest algorithm, an ensemble of particularly constructed decision trees, effectively overcomes this challenge. And we will focus on the Random Forest algorithm.

Random Forest

Random Forest (Breiman, 2001) algorithm in generating multiple decorrelated decision trees averages their predictions to yield a single prediction. Relying on the premise that averaging a set of independent observations having equal variances, this algorithm decreases the variance of the mean of the observations. The algorithm first generates a large number, say 'B,' bootstrapped samples from the training dataset. It then fits and trains the Decision Tree model on each of these B bootstrapped samples. The algorithm fits the decision trees on to the bootstrapped samples such that a random sample of 'm' features are considered as split candidates every time a split is made, rather than the entire set of features. Anytime a split is made, a fresh sample of random 'm' features are chosen for split consideration. Generally, 'm' is the square root of the total number of features. By drawing a fresh sample of 'm' features, the algorithm allows every feature to be considered for a split. This in turn produces uncorrelated decision

trees which result in uncorrelated predictions. Averaging these uncorrelated predictions leads in a reduction of the variance of the ensemble method.

More rigorously, if $\hat{f}^1(x), \hat{f}^2(x), ..., \hat{f}^B(x)$, are the predictions that go with the B distinct decorrelated decision trees for the test observation x, then the Random Forest offers the prediction,

$$\hat{f}_{RF}(x) = \frac{1}{B}\sum_{b=1}^{B} \hat{f}^b(x) \tag{25}$$

Note that the B decorrelated decision trees are grown deep and, therefore, register high variance and low bias. However, by averaging these decorrelated trees, the resulting Random Forest model achieves lower variance which improves its prediction accuracy.

#### 3.5.2.3 Support Vector Machines

Support Vector Machine (SVM) is a machine learning algorithm predominantly applied to binary classification problems. Its approach builds on the Maximal Margin Classifier algorithm applied in classifying linearly separable observations. Since most datasets cannot separate the observations by a linear boundary, the Maximal Margin Classifier has limited applications. By introducing Soft Margin and Kernel concepts to the Maximal Margin Classifier, SVM can classify observations with non-linear decision boundaries. Soft Margin is a boundary that basically classifies the observations into two different classes, though it cannot be said to do this perfectly. It misclassifies a few observations for the sake of improving its classification for a majority of training observations and achieving better robustness to individual observations. Further, to account for non-linear decision boundaries, SVM enlarges the feature space efficiently using specific functions called Kernels that quantify the level of similarity between the two observations. In adopting appropriate Soft Margin and Kernel, the resulting SVM model achieves lower variance and accounts for non-linear decision boundaries.

Maximal Margin Classifier relies on the existence of a hyperplane.[58] If a hyperplane exists, then this could act as a classifier such that an observation belonging to one side of the hyperplane is classified as class 1; if the observation belongs to the other side, then it is classified as class 2. Thus, an observation X belongs to class 1 if, say, for example,

$$f(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p > 0 \tag{26}$$

[58] A hyperplane is a linear boundary that separates a dataset's observations into two different classes. For instance, consider a two-dimensional feature space such that its observations could be separated by a linear boundary. In this case, the linear boundary, a hyperplane, is a line that divides the two-dimensional feature space into halves. Formally, $\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0$ is a hyperplane in a two-dimensional scenario where $\beta_0$,$\beta_1$, and $\beta_2$ are the parameters. The idea behind this notation could be extended to any arbitrary p-dimension feature space, where a hyperplane is defined as an affine subspace of dimension $p-1$. In other words, a hyperplane can be thought of as a flat subspace of dimension $p-1$ that divides the feature space into halves and follows the definition $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p = 0$.

And it belongs to class 2 if,

$$f(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p < 0 \tag{27}$$

Additionally, the magnitude $f(X)$ acts as a measure of confidence in the class assignment. If $f(X)$ is far from zero, then we can be confident about the class assignment. Whereas if f(X) is close to zero, then the class assignment may not be reliable.

Once we establish the existence of a hyperplane, then the Maximal Margin Classifier qualifies as the optimal hyperplane. Thus, it is the hyperplane that has the largest minimum distance from the training observations. We expect the optimal hyperplane to have the largest minimum distance from the training observations such that it can restore confidence in the class assignment of the observations. Once the Maximal Margin Classifier is located, the algorithm assigns a test observation to a class depending upon which side of Classifier it lies.

It so happens that the Maximal Margin Classifier depends only on a few training observations called the support vectors. Shifting a support vector or introducing a new observation that lies within the Margin of the optimal hyperplane could result in a new optimal hyperplane. This suggests that the algorithm is prone to overfitting the training dataset. To sidestep overfitting, we misclassify a few training observations for achieving better robustness to individual observations and assigning most of the training observations to the correct classes. More tolerant to a few misclassifications, the new classifier is called the Soft Margin Classifier. The number of misclassified observations violating the optimal hyperplane is governed by a tuning parameter. Much like the Maximal Margin Classifier, the Soft Margin also depends solely on the support vectors. The optimization problem for the Soft Margin Classifier could be modified by including additional functions to its features so that it could classify observations that could only be separated by a non-linear boundary. However, including additional functions could render the algorithm computationally expensive. Therefore, to obtain a computationally feasible non-linear decision boundary, SVM introduces Kernels to the Soft Margin Classifier.

#### 3.5.2.4 Gradient Boosted Models

A recently developed 'black-box' machine learning algorithm, Gradient Boosting Machine (GBM) has gained popularity for its high predictive accuracy. Being highly flexible, it could also be applied to a wide range of problems. GBM is an ensemble of weak predictive models where a weak model is defined as one whose prediction accuracy is only marginally better than random guessing. Any model can be a candidate for a weak model, however, for classification problems, such as ours, Classification Decision Trees are predominantly used (Kuhn et al., 2013). Our chosen weak predictive models, the Classification Decision Trees, are generally grown shallow with the number of splits ranging from 1-6. For our dataset, we note that GBM with 4 splits yield optimum results.

GBM was inspired by another boosting algorithm called AdaBoost, developed by Freund et al. (1996). In AdaBoost, a weak predictive model is fit to the weighted residuals of the ensemble created at the previous step so that the new weak predictive model could improve upon the errors made by the previous ensemble. In other words, a weak model is fit, in iteration, $i+1$, to the residuals of the ensemble created in iteration $i$, such that the residuals corresponding to the incorrectly predicted observations by the ensemble are assigned higher weights compared to those predicted correctly. Assigning higher weights to the observations whose response values are difficult to predict, allows the new weak model to focus on improving the prediction accuracy for these observations, hence improving the overall prediction accuracy for the whole ensemble.

Much like AdaBoost, GBM algorithm consists in fitting weak predictive models sequentially to the ensemble such that their inclusion improves the predictive performance of the whole ensemble. The weak predictive models are constructed such that these models and the negative gradient of the loss function associated with the whole ensemble are maximally correlated (Friedman, 2001). Below, we outline the GBM methodology.

Consider a training dataset $(\mathbf{x}_i, y_i)_{i=1}^N$ where $\mathbf{x}$ denotes the explanatory variables and $y$ denotes the response variable such that the true relationship between $\mathbf{x}$ and $y$ is given by $f$. We estimate a model $\hat{f}(\mathbf{x})$ such that it minimizes the expected value of the loss function[59] $L(y, f(\mathbf{x}))$,

$$\hat{f}(\mathbf{x}) = y \tag{28}$$

$$\hat{f}(\mathbf{x}) = argmin_{f(\mathbf{x})} E_{\mathbf{x}}[E_y(L(y, f(\mathbf{x})))|\mathbf{x}]$$

In restricting the search for the estimated model to the family of parametric functions, we consider the following "additive" expansion for the true function (in the equation below, M is the number of iterations),

$$f(\mathbf{x}; \{\beta_m, \mathbf{a}_m\}_1^M) = \sum_{m=1}^{M} \beta_m h(\mathbf{x}; \mathbf{a}_m) \tag{29}$$

In the above function, $h(\mathbf{x}; \mathbf{a})$ is a parameterized function of the explanatory variables $\mathbf{x}$, char-

[59] Since our response variable is binary, we consider the binomial loss function.

acterized by the parameters $\mathbf{a} = \{a_1, a_2, \dots\}$. In our case, $h(\mathbf{x}; \mathbf{a}_m)$ is a shallow classification tree and therefore the parameters $\mathbf{a}_m$ are the split variables, split locations, and the modes of the terminal node for the individual trees.

By choosing a parameterized model $f(\mathbf{x}; \mathbf{P})$, where $\mathbf{P} = \{P_1, P_2, ...\}$ is a finite set of parameters, the function optimization problem changes to the following parameter optimization problem,

$$\mathbf{P}^* = argmin_{\mathbf{P}} \Phi(\mathbf{P}) \tag{30}$$

where

$$\Phi(\mathbf{P}) = E_{y,\mathbf{x}} L(y, f(\mathbf{x}; \mathbf{P})) \tag{31}$$

We, therefore, get

$$\hat{f}(\mathbf{x}) = f(\mathbf{x}; \mathbf{P}^*) \tag{32}$$

Applying numerical optimization methods to solve for $\mathbf{P}^*$ imposes the solution for the parameters as $\mathbf{P}^* = \sum_{m=0}^{M} \mathbf{p}_m$. In this solution for $\mathbf{P}^*$, $\mathbf{p}_0$ and $\{\mathbf{p}_m\}_1^M$ are the initial guess and the successive increments ("boosts"), respectively. Each "boost" depends on the sequence of preceding "boosts" and to solve the optimization problem, the algorithm chooses Steepest-descent numerical minimization method. In defining the increments $\{\mathbf{p}_m\}_1^M$, first the gradient, $\mathbf{g}_m$, is computed,

$$\mathbf{g}_m = \{g_{jm}\} = \{[\frac{\partial \Phi(\mathbf{P})}{\partial P_j}]_{\mathbf{P}=\mathbf{P}_{m-1}}\} \tag{33}$$

where $\mathbf{P}_{m-1} = \sum_{i=0}^{m-1} \mathbf{p}_i$ . The increment is then defined as $\mathbf{p}_m = -\rho_m \mathbf{g}_m$, where,

$$\rho_m = argmin_{\rho} \Phi(\mathbf{P}_{m-1} - \rho \mathbf{g}_m) \tag{34}$$

In the above notation, $-\mathbf{g}_m$ is the direction of "steepest-descent" and $\rho_m$ is the "line search" along this direction.

Contrarily, we can also apply numerical optimization in the function space. In other words, we treat $f(x)$ as a parameter and minimize $\Phi(f) = E_{y,\mathbf{x}} L(y, f(\mathbf{x})) = E_{\mathbf{x}}[E_y(L(y, f(\mathbf{x})))\mathbf{x}]$.

We consider the solution to have the following functional form,

$$\hat{f}(\mathbf{x}) = \sum_{m=0}^{M} f_m^*(\mathbf{x}) \tag{35}$$

where $f_0^*(\mathbf{x})$ and $\{f_m^*(\mathbf{x})\}_1^M$ are the initial guess and increment functions ("boosts") defined by

the optimization, respectively. Each "boost" is updated as follows,

$$f_m^*(\mathbf{x}) = -\rho_m g_m(\mathbf{x}) \tag{36}$$

where

$$g_m(\mathbf{x}) = [\frac{\partial \phi(f(\mathbf{x}))}{\partial f(\mathbf{x})}]_{f(\mathbf{x})=f_{m-1}(\mathbf{x})} = [\frac{\partial E_y[L(y,f(\mathbf{x}))|\mathbf{x}]}{\partial f(\mathbf{x})}]_{f(\mathbf{x})=f_{m-1}(\mathbf{x})} \tag{37}$$

is the gradient[60] and

$$\rho_m = argmin_\rho E_{y,\mathbf{x}} L(y, f_{m-1}(\mathbf{x}) - \rho g_m(\mathbf{x})) \tag{38}$$

is the "line search" along the direction of $-g_m$. This non-parametric approach can no longer be applied when the joint distribution of $(\mathbf{x}, y)$ is estimated by the finite sample $(\mathbf{x}_i, y_i)_{i=1}^N$. To sidestep this, we can consider a parameterized form, as assumed in case of the parametric method discussed, thereby converting the optimization problem to a parametric optimization problem,

$$(\beta_m, \mathbf{a}_m)_1^M = argmin_{\{\beta_m' \mathbf{a}_m'\}_1^M} \sum_{i=1}^N L(y_i, \sum_{m=1}^M \beta_m' h(\mathbf{x}_i; \mathbf{a}_m')) \tag{39}$$

If the given approach also fails, then the "greedy stagewise" can be adopted as follows,

$$(\beta_m, \mathbf{a}_m) = argmin_{\beta,\mathbf{a}} \sum_{i=1}^N L(y_i, f_{m-1}(\mathbf{x}_i) + \beta h(\mathbf{x}_i; \mathbf{a})) \qquad For \quad m = 1, 2, \ldots, M \tag{40}$$

And the ensemble is updated as follows,

$$f_m(\mathbf{x}) = f_{m-1}(\mathbf{x}) + \beta_m h(\mathbf{x}; \mathbf{a}_m) \tag{41}$$

Thus, the choice of the loss function and weak predictive models determine the model properties of GBM. However, these choices in providing the algorithm with high flexibility render their applicability to a wide range of problems.

### 3.5.3 Model Evaluation

We now discuss the performance metrics used to evaluate our models. For evaluating the out-of-sample predictions, the data sample is split into training and test samples. The models are trained on the training sample and their predictive performance is estimated on the test sample, via its confusion matrix (Figure 1). A confusion matrix tabulates a model's class predictions against the actual class assignment of the observations. We label the entries of the confusion matrix as follows:

**TP:** the number of **true positives**, i.e., positive class observations that the model has correctly classified.

[60] $\phi(f(\mathbf{x})) = E_y[L(y, f(\mathbf{x}))|\mathbf{x}]$ and $f_{m-1}(\mathbf{x}) = \sum_{i=0}^{m-1} f_i^*(\mathbf{x})$

**TN:** the number of **true negatives**, i.e., negative class observations that the model has correctly classified.

**FN:** the number of **false negatives**, i.e., positive class observations that the model has incorrectly classified.

**FP:** the number of **false positives**, i.e., negative class observations that the model has incorrectly classified.

[ Please insert Figure 1 about here.]

We now define our metrics, true positive rate (TPR) and false positive rate (FPR), with reference to the confusion matrix.

TPR, also referred as sensitivity and recall, measures the proportion of positive observations correctly classified by a model:

$$TPR = \frac{TP}{(TP+FN)} \tag{42}$$

FPR, or fall-out, measures the proportion of negative observations misclassified by a model:

$$FPR = \frac{FP}{(FP+TN)} \tag{43}$$

Both TPR and FPR lie between 0 and 1. Typically, we want TPR to be as high as possible and FPR to be as low as possible. Unfortunately, these two metrics do not vary independently of each other, unless we deal with a perfect model. To achieve high TPR, we require a more sensitive model, though its inclusion would mean higher false positives, i.e., higher FPR. This trade-off is a general feature of any classification model.

Most ML classification algorithms estimate the probability of an observation belonging to the positive class. Typically, a value of 50% is used as the probability threshold, i.e., an observation whose estimated probability greater than the threshold is assigned the positive class; whereas, if the estimated probability is less than the threshold, it is assigned the negative class. Lowering the threshold increases the number of true positives, though it also increases the number of false positives. Raising the threshold lowers the number of false positives, though it comes at the expense of reducing the number of true positives. Therefore, to measure the overall performance of a model, we plot the receiver operator characteristic (ROC) curve. ROC curve is the graphical representation of the relationship between the true positive rate and false positive rate, when the probability threshold is varied.

Figure 2 shows a typical ROC curve for a classification model. Each point on the curve pro-

vides the TPR (y-coordinate) and FPR (x-coordinate) corresponding to a probability threshold. Ideally, a model with TPR equal to 1 and FPR equal to 0 yields the best predictive capacity. However, in practice, we choose a model that hugs the top left corner of the ROC curve. Additionally, to measure the model's out-of-sample predictive performance we compute the area under the ROC curve (AUC). AUC lies between 0 and 1. A model with AUC of 0.5 is no better than randomly guessing (random classifier) the class for an observation; a model with AUC less than 0.5 performs worse than the random classifier; and a model with AUC greater than 0.5 demonstrates predictive capacity.

[ Please insert Figure 2 about here.]

### 3.5.4 Predictor Importance

Finally, we investigate the relative importance of features in determining whether a transaction is fraudulent. In case of logistic regression, we use the statistical significance, and the magnitude of coefficient estimates to infer the relative importance of features. For random forests and gradient boosted machines, we estimate the total decrease in node purity corresponding to each predictor. Given that the SVM algorithm does not naturally extend itself towards estimating feature contribution, constructing a heuristic is in order. This method, unfortunately, does not provide consistent and reliable estimates. Therefore, we do not compute feature importance for the SVM model. We choose the models estimated on hybrid-sampled dataset to compute feature importance, since these models outperform the models fitted on datasets resampled by other techniques employed in this study.

## 3.6 Results

This section presents results of both our baseline empirical and robustness tests. We discuss the baseline results in subsection 3.6.1. The results of the robustness tests are discussed in subsections 3.6.2 and 3.6.3.

### 3.6.1 Model Performance and Interpretation

We first determine whether the various country-level features employed in our study can detect money-laundering at the financial institution. To meaningfully gauge the predictive capacity of these features, we first decompose our dataset into transactions involving private customers (people-related) and corporate clients (corporate-related).[61] We then train our models on the country-level attributes of the people-related, corporate-related, and combined dataset to estimate the out-of-sample performance of our models. We train 48 models; 4 machines learning algorithms trained on 3 datasets (people-related, corporate-related and the combined dataset) that are balanced by 4 balancing techniques. A randomized 50:50 split is performed on the

[61] See Table 1 and Table A1 (Internet Appendix A).

datasets to create training and test datasets.[62] We further perform cross-validation to test the validity of our models and estimate relative importance of various country-level features.

#### 3.6.1.1 Predictive capacity of Country-level features

Table 3 shows the TPR, FPR, and AUC results of the models trained on the country-level features of the three datasets. Our features include the Hofstede country-specific culture dimensions and two institution quality indices for the customer's country of residence and origin/destination country of the wire.[63] These features are: $CPI_R$, $CPI_W$, $FSI_R$, $FSI_W$, $IDV_R$, $IDV_W$, $MAS_R$, $MAS_W$, $PDI_R$, $PDI_W$, $UAI_R$, and $UAI_W$. For models trained on the combined dataset, we note that the AUCs are in the 0.70-0.80 range. This demonstrates that our models can discern between suspicious and legitimate transactions. We find that our models can discern better for the corporate-related dataset with AUCs as high as 0.88. We further find evidence for predictive capacity for the country-level features for the people-related data. However, compared to the combined and corporate-related data, these results are modest with AUCs in the 0.65-0.72 range.[64]

We further note that all the models trained on datasets balanced by the hybrid-sampling technique consistently provide significant out-of-sample performance. Additionally, we find that the RF and GBM models have the best out-of-sample performance for all the three datasets balanced by the under- and hybrid-sampling techniques.

[Please insert Table 3 about here.]

#### 3.6.1.2 Determining the validity of our models using cross-validation techniques

To determine the validity of our models we perform K-fold cross-validations. K-fold cross-validation estimates how well a model generalizes an independent dataset by dividing the dataset into K equal parts, using one part as a hold-out test set, and training the model on the remaining K-1 parts. This is then repeated K times, such that each of the K equal parts is considered for a test dataset. The out-of-sample model performance is then computed as the average of the K results. We perform 5- and 10-fold cross-validations, and for each of the K instances, we train our models on 80% and 90% of the datasets, balanced by the hybrid-sampling method, respectively.[65] We then estimate the out-of-sample performance on the remaining 20%

[62] Except in the case of cross validation, where 80:20 and 90:10 splits are performed.

[63] Please see Table 2 for concise definitions.

[64] We further train our models on the Hofstede country-specific culture dimensions, excluding the institution quality indices. We report the models that include only the national culture dimensions are comparable to models including the institution quality indices as well. Please see Tables B1-B3 of the Internet Appendix B.These may reflect that national culture and national governance are endogenously related. However, our objective is to determine whether national culture is an effective predictor. We do not aim to identify a causal relationship between national culture traits and malfeasance in banks.

[65] We train our models on the three datasets balanced by the hybrid-sampling method since this method results in models with high predictive performance across all the three datasets.

and 10% of the datasets. In Table 4 we report the AUC metric, estimated by the cross-validation technique, to measure performance for all the models. The results demonstrate that the predictive capacity of country-level variables remain similar to that reported in Table 3. The low standard deviation ($\sigma$) further attests to the reliability of our models.

[Please insert Table 4 about here.]

#### 3.6.1.3 Investigating the relative importance of country-level features in detecting money-laundering

Table 5 presents the relative importance of our country-level features for the models trained on the three datasets.[66] We find that for both corporate-related and combined alerts, the individuality rating of both the customer's residence country ($IDV_R$) and country of wire origination/destination ($IDV_W$) are of paramount importance. This is followed by the corruption perception score of the country of wire origination/destination ($CPI_W$) and the customer's residence country ($CPI_R$) for the corporate-related alerts; and ($CPI_W$) and the financial secrecy score of the customer's resident country ($FSI_R$) for the combined alerts. For people-related alerts, the corruption perception score for the country of wire origination/destination ($CPI_W$) and the financial secrecy score of the resident country ($FSI_R$) are the two most important features, followed by the $CPI_R$ and $IDV_R$.

[Please insert Table 5 about here.]

### 3.6.2 Can we improve the predictive capacity of our models by enlarging the feature space?

In this section, we extend our feature space to include account- and transaction-level variables. We further include the proprietorial risk score (PROP Score) in our enlarged feature space to assess the predictive capacity of our models.[67]

#### 3.6.2.1 Predictive capacity of country-, account-, and transaction-level features in detecting money-laundering

We further extend our feature space to include customers' account- and transaction-level information to determine whether we could improve the predictive capacity of our models. This extends our feature space to include 24 features with 12 country-level features, 4 account-level features, and 8 transaction-level features.[68]

[66] We do not report feature importance results for the SVM model since there does not exist a reliable model-specific feature importance method for SVM algorithm.

[67] All features are defined in Table 2.

[68] We include an additional feature, Customer Age, in the people-related models which extends the feature set to include 25 predictors.

Table 6 presents the TPR, FPR, and AUC scores for models trained on the enlarged feature space. These models enhance the predictive capacity across all the models reported in Table 3, with AUCs ranging between 0.72-0.91, 0.83-0.94, and 0.60-0.85, on the combined, corporate- and people-related datasets, respectively. We further note that the models trained on the datasets balanced by the hybrid technique are better able to discern between a fraudulent and non-fraudulent transaction with AUC scores between 0.75-0.91, 0.85-0.94, and 0.71-0.85 for the combined, corporate-, and people-related datasets, respectively. We report a significant increase in the predictive capacity of our models across all the three datasets. We again find that the RF and GBM models with under- and hybrid-sampling are the optimal models.

[Please insert Table 6 about here.]

#### 3.6.2.2 Predictive capacity of country-, account-, and transaction-level features along with the proprietorial risk score in detecting money-laundering

Finally, we include the proprietorial risk score, PROP Score, to our enlarged feature space to determine whether its inclusion markedly enhances the predictive capacity of the models reported in Table 6.

We report the out-of-sample performance of these models in Table 7. Interestingly, we find only a slight improvement, of approximately 1-2% on average, in performance. This indicates that models with the country-, account- and transaction-level information provide useful predictive power.

[Please insert Table 7 about here.]

### 3.6.3 Does national culture traits remain useful in the extended dataset?

In this section, we investigate whether the country-specific culture and institution quality indices pertaining to customer's residence country and the country of wire origination/destination remain useful in detecting money-laundering in the enlarged feature space.

#### 3.6.3.1 Does national culture traits remain useful in comparison with account-level and transaction-level variables?

We estimate feature importance for models reported in Table 6 to determine whether country-level features of the customers provide useful predictive capacity in detecting fraudulent wire transactions in the enlarged feature space. We present these results in Table 8. We note that for corporate-related alerts, the county-level features that rank among the top five features are the individuality rating of the customer's country of residence ($IDV_R$), individuality rating of the country of wire origination/destination ($IDV_W$), and the uncertainty avoidance cultural trait of the customer's residence country ($UAI_R$). We further find that the power-distance

index score of the customer's residence country ($PDI_R$) informs the customer's predilections for committing financial misconduct. For people-related alerts, the individuality score of the customer's residence country ($IDV_R$), corruption perception score of the country of wire origination/destination ($CPI_W$), and financial secrecy score of the customer's country of residence ($FSI_R$) are the most important county-level features that rank among the top ten features. These results provide evidence of the usefulness of culture traits of customers for detecting both corporate and individual malfeasance. However, the country-level features are more pronounced in detecting corporate malfeasance than individual malfeasance. For the combined alerts, we note that $IDV_R$, $IDV_W$, and $FSI_R$ rank among the top ten features. This further provides evidence of the usefulness of the country-level features in detecting malfeasance.

[Please insert Table 8 about here.]

#### 3.6.3.2 Does national culture traits remain useful in comparison with a proprietorial risk score along with account- and transaction-level features?

Table 9 reports the feature importance for models reported in Table 7. For corporate-related alerts, we again find that the individuality scores of both the country of the wire origination/destination ($IDV_W$) and customer's resident country ($IDV_R$) are important country-level features. These features also rank among the top five features influencing a customer's predilections for committing financial misconduct. We further note that the corruption perception score of the country of wire origination/destination ($CPI_W$) and power-distance index score of the customer's residence country ($PDI_R$) are among the top ten features. Interestingly, we find that $IDV_W$, $IDV_R$, and $CPI_W$ have higher predictive capacity than the proprietorial risk score. However, in case of people-related alerts, the PROP Score is the most important feature. This suggests that the proprietary algorithm, used by the financial institution, is more effective in detecting fraudulent transactions pertaining to individual accounts than for corporate accounts. Further, in case of people-related alerts, the financial secrecy score of the customer's residence country ($FSI_R$), corruption perception score of the customer's residence country ($CPI_R$), and corruption perception score of the country of wire origination/destination ($CPI_W$) rank among the top ten features in detecting money-laundering in our models. For the combined alerts, the features that influence the models in decreasing order are $IDV_W$, *PROP Score*, $IDV_R$, and $FSI_R$. These features also rank among the top ten features. In addition to results reported in Table 8, these results further underline the usefulness of adopting country-specific features to complement current account and transaction variables for AML monitoring.

[Please insert Table 9 about here.]

## 3.7 Discussion and Ethical Framework

Since Donaldson and Dunfee (1994), scholars have subscribed to the notion that business ethics research can either be informed by empirical ideas or normative concepts and prescriptive ideas. Business ethics research informed by normative concepts, although not necessarily accessible through empirical analysis, suggests what societies should do. This line of research proves that empirical analysis is often not the appropriate tool to determine what societies "ought" to do (see also Sorley (1885)). Thus, money-laundering detected through out-of-sample predictive accuracy alone, as this paper establishes, inadequately explains why a machine learning alert model should be deployed.

The potential for AI applications' unethical repercussions, especially those that impact people's wellbeing, is immense. Examples include recruitment, promotion, flight risk, and cessation of employment algorithms as well as credit extension, insurance risk scoring, and dynamic pricing algorithms, among others. Fraud detection, mediated through machine learning, arguably falls on the lower end of the spectrum of potentially unethical AI, considering its goal of mitigating financial malfeasance. Nevertheless, it is critically important to consider the ethical implications of factoring in nationality as a prompt for scrutinizing individuals.

With a view, hence, to "giving voice to values" (Arce and Gentile, 2015), we seek to identify the ethical implications of incorporating profiling, never mind whether it is intentional or not, within machine learning algorithms. We discuss ethical issues pertinent to the deployment of national culture in machine learning in general and money-laundering alert models in particular.

We frame our discussion around central ethical questions, including 1). Do public good concerns in countering money-laundering outweigh 'collective treatment' in algorithmic national profiling? 2). Do those issuing the alerts permitted to avail of the personal data? 3). Who is responsible for the design of an algorithm? 4). Are algorithms accountable? 5). Are the algorithms used for detection or prediction? And are there subtle distinctions between them? 6). Do alert models reflect global, national or sub-national, public or private regulation? 7). Do the existing algorithms exacerbate tangential social biases?

### 3.7.1 Do public good concerns in countering money-laundering outweigh 'collective treatment' in algorithmic national profiling?

Alter and Darley (2009) define collective treatment as 'the act of behaving toward more than one individual uniformly.' Contrastingly, in individualized treatment, individuals are treated differently based on certain criteria. An example of collective treatment is punishing a group for being offenders as opposed to prosecuting its individual members commensurate with their level of crime. According to Alter and Darley (2009), collective treatment is predicated on a group's shared salient features (such as race and ethnicity, among others) that are used to

stereotype them. Thus, people belonging to the same group are treated as interchangeable members of the group. As noted by Brewer and Harasty (1996), Campbell (1958), Dasgupta et al. (1999) among others, such salient features can include race, ethnicity, socioeconomic status, religion, physical appearance, nationality, and even debility. Clearly, national culture can also be featured to characterize and homogenize a people, especially in alert models that excavate cultural factors to detect fraud. However, the chief danger of collective treatment lies in the prospect of individuals in positions of authority administering it to reward, punish, or restrict the rights of a group within a population. For instance, a judge who sentences a criminal gang rather than its individuals etc. One advantage of machine-learning-based alert models is in sidestepping arbitrary individual choices to impose or not to impose collective treatment.

### 3.7.2 Do those issuing the alerts permitted to avail of the personal data?

The legality of gathering and mining certain data does not in itself make it ethical. After all, 'Ethics' constitutes a set of moral codes beyond legally stipulated minimums. As regards mining of data inhering in machine learning procedures, questions of ethicality invariably arise.

Whether the institution conducting machine learning is allowed to use the data in its algorithms is at once an important legal and ethical issue. While there may be legal barriers to using particular data, ethical issues too loom over issues of legality. In many cases, machine learning may end up employing data without proper permissions. As noted by Adomavicius and Tuzhilin (2001), data mining in the context of individuals is viewed as either 'factual' viz. who the customer or 'transactional', what the customer has done or is doing (see also Cook (2008)). Adomavicius and Tuzhilin (2001) suggest the latter is more commonly used for identifying a criminal as well as more commonly contested for intruding on individual privacy. However, the money-laundering alert model of this paper suggests that simply using data about who the customer is, i.e., the customer's home country, can generate area-under-the-curve predictions that are almost 90 percent successful. So, using national culture as a predictor brings the potential advantage of circumventing intrusive gathering of customer behavior. Use of national culture avoids issues of employing personal data. The sweeping aggregate generality of national culture helps avoid invasive use, more often than not without permission, of individual characteristics. Thus, in respect of including national culture in machine-learning models, a relevant question arises, if not national factors, then what other factors? And what would be the alternative set of implications? Overall, as regards permissions to use data, national culture, while arguably a rough profiling of people from respective nations, avoids the use of more individual and likely personal data. In effect, an ethical minefield opens up in the interstices of national culture profiling's sweeping generalities, and the potential invasion of privacy in the collection of personal, often transactional, data.

### 3.7.3 Who is responsible for algorithmic design?

Martin (2019) compellingly argues why developers of algorithms should be held accountable for the unethical consequences of their productions. Asserting that inscrutable algorithms necessitate algorithmic accountability, he wants to fix the responsibility on the developer for the unethical consequences of her work. He also illustrates numerous examples of disturbing and unintended moral consequences of algorithms. In silently structuring lives, algorithms can create bespoke pricing for online products to individuals (Angwin et al., 2016), determine if an applicant's loan will be granted (Kharif, 2016), show specific online content to influence the decision of voters during the presidential elections (O'Neil, O'Neil), determine if parole will be granted to an incarcerated inmate (Angwin et al., 2016; Wexler, 2017), among many others.

It is also worth considering whether academic authors should be held responsible for their ideas and findings. This seems stretching the argument somewhat because that would amount to inhibiting scholarly investigation. However, as authors, we are concerned that our paper's evidence that national factors, particularly national culture, might be particularly efficacious in money-laundering alert models, may in itself have various moral consequences. As Martin (2019) aptly points out, algorithms are inherently value-laden and need to be built to preserve the stakeholder's "rights and dignity."

### 3.7.4 Are algorithms accountable?

Another concern of machine learning is algorithmic accountability (Buhmann et al., 2019; Martin, 2019; Seele et al., 2019). This includes how algorithms are established, if their hypotheses are either explicit or implicit. However, from an alternative perspective, such transparency would also provide criminals access to the factors used in respective algorithms, enabling subsequent avenues of evasion. However, with respect to national culture, can knowledge of national culture's inclusion in an alert algorithm be gamed by would-be illicit actors? Would this encourage actors to channel banking transactions through other countries with differently identified cultural characteristics? Do clever money-launderers already have a cue that culture is a criterion to establish money-laundering alerts? Overall, being transparent about the use of national culture in machine-learning algorithms is perhaps less deleterious to the stakeholders than transparency about other details of algorithms.

### 3.7.5 Are the algorithms used for detection or prediction? And are there subtle distinctions between them?

Another fundamental distinction which touches on the ethics of machine learning algorithms concerns whether alert procedures are to be used in detection or prediction. This also involves broader issues of the implications of ex-ante or ex-post investigation. Depending on the context, prediction can lead, or not lead, to particular consequential actions. For instance, an algorithm

to predict personal loan default might lead to denial of financing to a worthy applicant (Fuster et al., 2018). It is also possible that a prediction algorithm to identify possible bank fraud might facilitate time and resources being devoted to rigorous scrutiny. In this regard, using machine learning to detect money-laundering, as in the example of this paper, could be viewed, as simply reducing the costs of detection, rather than establishing an unfair barrier to banking inclusion.

In contrast, using machine learning to predict what might take place creates identifiable issues of fairness. For instance, highly controversial practice obtains in the US of using ethnic and racial profiling to predict whether prison inmates under parole consideration will recidivate (Hartney, 2009). This has even been extended to machine learning algorithms (Berk, 2017; Lee, 2018). Examples such as this display an obvious unfairness and social injustice. This is inherently different from identifying whether money-laundering has already occurred.

Or is it so different? Certainly, there is the possibility of organizations transferring usage of algorithms from detection to pre-emption. In which case, factors included in detection are now used to unfairly exclude. Further, identifying persons from particular countries in the context of global regulation appears at least somewhat differently from law enforcement in a particular country or sub-national component of a country targeting certain citizens based on demographic characteristics for additional scrutiny. Or is a case of global regulation being that different? Certainly, this issue beckons much further reflection.

Of additional concern, the distinction between detection and prediction gets blurred for situations where 'everyone' is engaged in a particular illegal action. For instance, it is not uncommon on many of the US interstate highways, where speed cameras are generally not used as widely as in other countries, for the great majority of drivers to be exceeding the speed limit. However, it is generally the case in the US that African American drivers are much more likely to be pulled over by the police (Harris, 1996). This could, of course, be due to racial prejudice. But it is commonly believed that speeding African American drivers provide greater opportunities for law enforcement to detect their other violations because they have a higher percentage of criminal records. This fuels a self-fulfilling prophecy. Another example is the controversial practice of the US Internal Revenue Service (IRS) closely scrutinizing tax-exempt status of vocal anti-tax groups. On the one hand, this practice could be viewed as targeting opposition to the government. On the other hand, is it not reasonable to consider vocal anti-tax groups as more likely to evade taxes? Detaining African Americans for traffic violations has obvious agendas of intimidation and social repression. Greater scrutiny of the taxes of groups with a particular political orientation presents similar concerns.

### 3.7.6 Do alert models reflect global, national or sub-national, public or private regulation?

Global regulation that focuses more on some countries than others arguably presents a different context than inflection of legal authority unevenly within a particular country or sub-national jurisdiction. This is perhaps because global monitoring, as with money-laundering alerts, is often in the realm of the private sector. Consequently, much of the public sector's social inequities that cause the isolation of groups within a society is avoided. However, if we consider the world as a globalized society, then such distinctions diminish. It is arguable that the governance mandate of private global firms needs to be evenly administered. An interesting parallel is the openly disclosed pillar of the microfinance industry to focus on women borrowers. In other words, the industry exhibits less inclination to grant loans to men (Aggarwal et al., 2015). Promoted as micro-finance to women that will offer great social outreach benefits, it is premised on the logic that women are more likely to pay back the loans availed from the industry. Perhaps identifying a particular gender as more likely to repay a loan is not fundamentally different from identifying people of particular national cultures as more likely to repay loans—or, more or less likely to indulge in money-laundering.

Issues of social fairness become much more glaring, however, when we consider the wide variety of research that seeks to model national at a sub-national level with demographic, particularly data on religion. For instance, Baxamusa and Jalal (2014) model religion as indicating levels of what national culture would describe as uncertainty avoidance. A significant problem with using national culture in algorithms consists in the potential for biases about particular national cultures to diffuse to sub-national levels.

### 3.7.7 Do the existing algorithms exacerbate tangential societal biases?

Another concern is whether the respective machine-learning algorithm in incorporating factors tangentially engenders implicit biases. Williams et al. (2010) provides an example of this. They highlight a case of poor verbal skills being correlated with higher incidence of scholastic cheating. Clearly, such reasoning can foster serious bias against immigrant communities or others with generally suboptimal skills in the given language of instruction. Or can establish bias against those with speech impediments, for instance. Do algorithms that incorporate national culture foster prejudice? Our study in this paper, for instance, suggests that people from individualist countries are more likely to engage in money-laundering. Consider the roles of profiling by national culture in other contexts. For instance, profiling potential CEOs or board members as to whether they would be effective in CSR advocacy, or whether they have optimal demographic characteristics (Johnson et al., 2013). Would certain CEO aspirants be disfavored or disqualified because of their country of origin or ethnic background?

The use of national culture in bank alert models unmistakably provides some predictive accu-

racy, even while it gives rise to a number of important social and ethical issues. We hope this paper would invite further analysis and discussion on this important issue.

## 3.8 Conclusion

Recent high-profile scandals involving banks such as Danske and Swedbank call for an urgent need to innovate in current AML protocols. The existing compliance requirements impose significant costs on financial institutions with negligible returns. Further, the current AML surveillance practices are painstakingly inefficient, time-consuming, and labor-intensive. Our paper examines the utility of incorporating national culture profiling in bank-level machine-learning informed alert models to detect money-laundering at a globally important financial institution.

Prior studies associate national culture dimensions with financial misconduct (Liu, 2016; De-Backer et al., 2015; Bame-Aldred et al., 2013); quality of ethical behaviour and perception (Armstrong, 1996; Davis and Ruhe, 2003; Getz and Volkema, 2001; Vitell et al., 1993; Volkema, 2004); and finance-related behaviour (Chui et al., 2010; Lievenbrück and Schmid, 2014; Aggarwal and Goodell, 2009; Chui et al., 2002; Shao et al., 2013; Aggarwal and Goodell, 2013). Given the breadth of scholarship that investigates national culture in the context of business behaviour, it stands to reason to study whether national culture impacts an individual's or corporation's predilection for bank fraud.

We test to establish how the utility of national culture traits informing a machine learning alert model helps in detecting money-laundering at a globally prominent financial institution. Pervasive across borders and undermining local economies, money-laundering remains an issue of global concern. In generating and disbursing illicit proceeds from criminal activities that have integrated into the financial system, money-laundering paves the way for further financial illegal activity. Using the financial institution's dataset of over 200,000 international wire transactions collected over a ten-year period, we built machine learning models that reference the levels of corruption and financial secrecy in a country, and the cultural measures of individualism, masculinity, power distance, and uncertainty avoidance. We find that besides the industry standard account- and transaction-level variables, the country-level variables significantly improve our models' predictive power, particularly in the category of corporate accounts. Using the machine learning algorithms to estimate the relative importance of the predictors in the most successful models involving corporate accounts, we discover that individualism scores for the customer's resident country and for the wire's country of origin/destination respectively, is by far the most important of the country-level variable and, indeed, of all the variables at work. As regards personal account models, the corruption perception score for the wire's country of origin/destination and the financial secrecy score for the customer's resident country prove to be the most important country-level variables, with other predictors outside of these proving

significant in predicting the incidence of suspicious wire activity. However, our results suggest that country-level data, particularly national culture scores of either the sender or receiver of wire transfers, either alone, or in combination with measures of corruption control and financial secrecy, provide highly effective prediction modelling. Given the social implications long identified with 'collective treatment,' our results provoke considerable reflection on the ethical concerns attendant on using country-level variables by financial institutions to fabricate money-laundering alert models.

Our findings indicate the importance of cultural and behavioural measures in assessing the potential for money-laundering and fraud in international money movement vis-à-vis corporate activity and provide strongly predictive models for detecting such behaviour. Furthermore, the models when applied to the segregated data sample (corporate account vs. individual account) demonstrate distinct differences in terms of predictive performance as well as feature importance. Practitioners can benefit by carefully configuring sample segmentation as well as feature selection. Applying a more contextual lens to current Anti Money Laundering (AML) surveillance practices may prove a valuable resource in the worldwide fight against money-laundering and fraud.

With a view, hence, to "giving voice to values" (Arce and Gentile, 2015), we seek to identify the ethical implications of incorporating profiling, never mind whether it is intentional or not, within machine learning algorithms. We discuss ethical issues pertinent to the deployment of national culture in machine learning in general and money-laundering alert models in particular. In this regard, we discuss central ethical questions, including 1). Do public good concerns in countering money-laundering outweigh 'collective treatment' in algorithmic national profiling? 2). Do those issuing the alerts permitted to avail of the personal data? 3). Who is responsible for the design of an algorithm? 4). Are algorithms accountable? 5). Are the algorithms used for detection or prediction? And are there subtle distinctions between them? 6). Do alert models reflect global, national, or sub-national, public, or private regulation? 7). Do the existing algorithms exacerbate tangential social biases?

## 3.9 Tables and Figures

**Figure 1: Confusion Matrix**

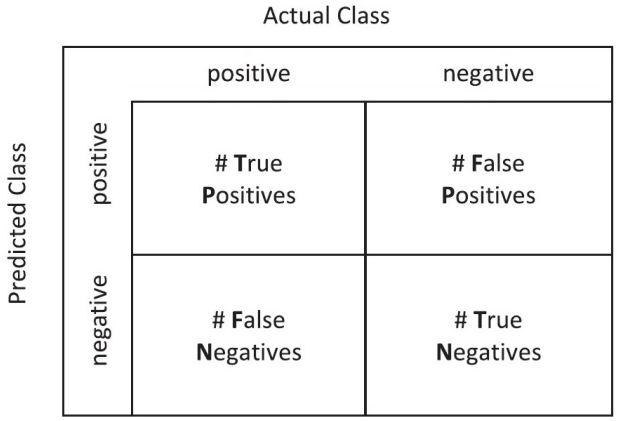



**Figure 2: ROC Curve**

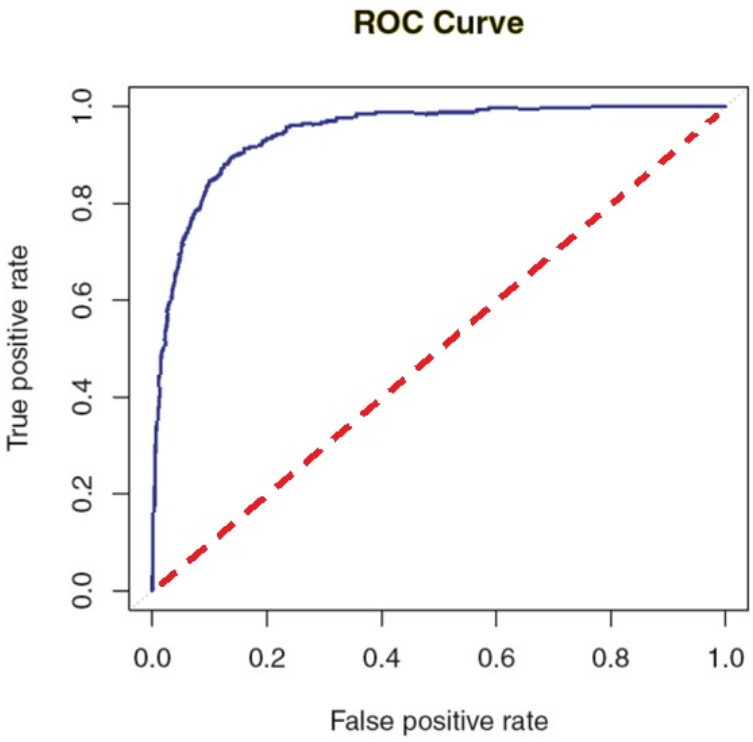


**Table 1: Data Cross-section and Sample Selection**

**Panel A: Alerts and Issue Cases by Year**

| | Combined | | Corporate | | People | |
|---|---|---|---|---|---|---|
| **Year** | **#Alerts** | **#Issues** | **#Alerts** | **#Issues** | **#Alerts** | **#Issues** |
| 2009 | 22,183 | 878 | 5,752 | 448 | 16,431 | 430 |
| 2010 | 23,154 | 485 | 6,643 | 215 | 16,511 | 270 |
| 2011 | 20,335 | 216 | 6,193 | 68 | 14,142 | 148 |
| 2012 | 18,572 | 143 | 5,298 | 29 | 13,274 | 114 |
| 2013 | 21,088 | 205 | 5,984 | 71 | 15,104 | 134 |
| 2014 | 11,098 | 87 | 2,617 | 41 | 8,481 | 46 |
| 2015 | 11,468 | 34 | 2,937 | 7 | 8,531 | 27 |
| 2016 | 11,779 | 71 | 2,841 | 14 | 8,938 | 57 |
| 2017 | 9,885 | 76 | 2,771 | 19 | 7,114 | 57 |
| 2018 | 4,355 | 11 | 1,236 | 2 | 3,119 | 9 |
| **Total** | **153,917** | **2,206** | **42,272** | **914** | **111,645** | **1,292** |

**Panel B: Sample Selection**

| | Combined | | Corporate | | People | |
|---|---|---|---|---|---|---|
| **Selection Criteria** | **#Alerts** | **#Issues** | **#Alerts** | **#Issues** | **#Alerts** | **#Issues** |
| All Alerts | 206,751 | 2,440 | 42,272 | 914 | 111,645 | 1,292 |
| Corp & Ppl Accounts | 153,917 | 2,206 | 42,272 | 914 | 111,645 | 1,292 |
| Country-level Variables | 74,832 | 1,183 | 30,303 | 524 | 44,529 | 659 |
| Account/Transaction-level Variables | 74,246 | 1,172 | 30,292 | 524 | 43,954 | 648 |

**Notes:** The table reports the cross-section of our data (Panel A) and the sample selection (Panel B). An alert is raised when a customer's wire activity raises certain flags and an Issue case indicates that the subsequent investigation has deemed the activity to be highly suspicious. The sample selection shows the number of alerts available our data set according to each criterion, applied in sequence. A more detailed description of our variables is available in Table 2.

## Table 2: Predictor Details

| Predictor Abbreviation | Details |
|---|---|
| **COUNTRY-LEVEL** | |
| **Corruption Perception Index** | Score of customer's residence country ($CPI_R$) and country of origin/destination of wire ($CPI_W$) according to Transparency International's Corruption Perception Index. |
| **Financial Secrecy Index** | Score of customer's residence country ($FSI_R$) and country of origin/destination of wire ($FSI_W$) according to Transparency International's Financial Secrecy Index. |
| **Individualism Index** | Score of customer's residence country ($IDV_R$) and country of origin/destination of wire ($IDV_W$) based on Hofstede's "Individualism" dimension of culture. |
| **Masculinity Index** | Score of customer's residence country ($MAS_R$) and country of origin/destination of wire ($MAS_W$) based on Hofstede's "Masculinity" dimension of culture. |
| **Power-Distance Index** | Score of customer's residence country ($PDI_R$) and country of origin/destination of wire ($PDI_W$) based on Hofstede's "Power-Distance" dimension of culture. |
| **Uncertainty Avoidance Index** | Score of customer's residence country ($UAI_R$) and country of origin/destination of wire ($UAI_W$) based on Hofstede's "Uncertainty Avoidance" dimension of culture. |
| **ACCOUNT-LEVEL** | |
| **Customer Age** | Age of customer associated with alert, at time of alert (CUS_AGE). |
| **Account Age** | Age of account associated with alert, at time of alert (ACC_AGE). |
| **Customer Net Worth** | Net Worth of customer associated with alert (NET_WRTH). |
| **Alert Supplier Code** | Code denoting source of alert, whether alert is generated by Business or Retail transactions (SUPP_CO). |
| **TRANSACTION-LEVEL** | |
| **Amount Transfers In** | Aggregate amount of incoming wire and electronic transfers over 180 days before alert ($TFI_{180}$). |
| **No. Transfers In** | Number of incoming wire and electronic transfers over 180 days before alert (#$TFI_{180}$). |
| **Amount Transfers Out** | Aggregate amount of outgoing wire and electronic transfers over 180 days before alert ($TFO_{180}$). |
| **No. Transfers Out** | Number of outgoing wire and electronic transfers over 180 days before alert (#$TFO_{180}$). |
| **Amount Checks In** | Aggregate amount of incoming checks over 180 days before alert ($CKI_{180}$). |
| **No. Checks In** | Number of incoming checks over 180 days before alert (#$CKI_{180}$). |
| **Amount Checks Out** | Aggregate amount of outgoing checks over 180 days before alert (($CKO_{180}$)). |
| **No. Checks Out** | Number of outgoing checks over 180 days before alert (#$CKO_{180}$). |
| **Proprietary** | |
| **PROP Score** | Risk score based on proprietary alert algorithm of financial institution. |

**Notes:** The table reports the complete set of predictors used in our models along with their definitions and abbreviations for reference. The "Wire" variables refer only to the wire transactions on the day of an alert whereas the "Transfer" and "Check" variables refer to all relevant transactions appearing on accounts associated with an alert in the 180 day period preceding that alert.

**Table 3: Country-level Models**

| | | Combined | | | Corporate | | | People | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Model** | **Balancing** | **TPR** | **FPR** | **AUC** | **TPR** | **FPR** | **AUC** | **TPR** | **FPR** | **AUC** |
| **LR** | **No Balancing** | 0.70 | 0.43 | 0.722 | 0.89 | 0.50 | 0.845 | 0.59 | 0.42 | 0.670 |
| | **Under-sampling** | 0.76 | 0.49 | 0.727 | 0.90 | 0.51 | 0.850 | 0.61 | 0.43 | 0.664 |
| | **Hybrid-sampling** | 0.76 | 0.50 | 0.726 | 0.90 | 0.47 | 0.851 | 0.58 | 0.40 | 0.670 |
| | **Synthetic-sampling** | 0.71 | 0.43 | 0.723 | 0.92 | 0.53 | 0.861 | 0.60 | 0.41 | 0.659 |
| **RF** | **No Balancing** | 1.00 | 1.00 | 0.543 | 1.00 | 1.00 | 0.674 | 1.00 | 1.00 | 0.505 |
| | **Under-sampling** | 0.71 | 0.40 | 0.741 | 0.89 | 0.41 | 0.875 | 0.66 | 0.41 | 0.702 |
| | **Hybrid-sampling** | 0.65 | 0.31 | 0.726 | 0.88 | 0.34 | 0.878 | 0.66 | 0.41 | 0.695 |
| | **Synthetic-sampling** | 1.00 | 1.00 | 0.696 | 1.00 | 1.00 | 0.859 | 1.00 | 1.00 | 0.641 |
| **SVM** | **No Balancing** | 0.53 | 0.47 | 0.521 | 0.42 | 0.42 | 0.504 | 0.62 | 0.56 | 0.516 |
| | **Under-sampling** | 0.78 | 0.60 | 0.704 | 0.88 | 0.59 | 0.805 | 0.66 | 0.44 | 0.660 |
| | **Hybrid-sampling** | 0.68 | 0.50 | 0.662 | 0.88 | 0.59 | 0.807 | 0.68 | 0.47 | 0.636 |
| | **Synthetic-sampling** | 0.77 | 0.51 | 0.645 | 0.89 | 0.60 | 0.816 | 0.59 | 0.41 | 0.610 |
| **GBM** | **No Balancing** | 0.87 | 0.60 | 0.768 | 0.91 | 0.49 | 0.878 | 0.84 | 0.56 | 0.719 |
| | **Under-sampling** | 0.87 | 0.59 | 0.770 | 0.85 | 0.41 | 0.870 | 0.83 | 0.57 | 0.708 |
| | **Hybrid-sampling** | 0.74 | 0.40 | 0.771 | 0.90 | 0.43 | 0.881 | 0.83 | 0.55 | 0.716 |
| | **Synthetic-sampling** | 0.68 | 0.41 | 0.724 | 0.94 | 0.60 | 0.868 | 0.73 | 0.57 | 0.660 |

**Notes:** The table reports the performance of our Country-level model using logistic regression (LR), random forest (RF), support vector machine (SVM) and gradient boosting (GBM) in combination with no balancing, under-sampling, hybrid-sampling and synthetic-sampling, respectively. The performance is measured using True Positive Rate (TP Rate), False Positive Rate (FP Rate) and Area under the ROC Curve (AUC). The data sample comprises of 74,724 alerts (30,292 corporate-related and 43,954 people-related) with 1,183 Issue cases (524 corporate-related and 648 people-related). The model has 12 predictors.

## Table 4: Cross-validation for Country-level Models with Hybrid-sampling.

| Panel A: 5-Fold Cross-validation on AUC scores | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Combined | | | | Corporate | | | | People | | | |
| **Round** | **LR** | **RF** | **SVM** | **GBM** | **LR** | **RF** | **SVM** | **GBM** | **LR** | **RF** | **SVM** | **GBM** |
| 1 | 0.717 | 0.722 | 0.656 | 0.765 | 0.758 | 0.774 | 0.785 | 0.813 | 0.658 | 0.662 | 0.594 | 0.683 |
| 2 | 0.726 | 0.726 | 0.705 | 0.777 | 0.856 | 0.862 | 0.811 | 0.896 | 0.674 | 0.706 | 0.663 | 0.724 |
| 3 | 0.737 | 0.729 | 0.705 | 0.766 | 0.861 | 0.873 | 0.829 | 0.902 | 0.671 | 0.675 | 0.598 | 0.709 |
| 4 | 0.743 | 0.739 | 0.678 | 0.789 | 0.852 | 0.872 | 0.842 | 0.887 | 0.660 | 0.681 | 0.606 | 0.723 |
| 5 | 0.762 | 0.768 | 0.725 | 0.797 | 0.820 | 0.848 | 0.833 | 0.874 | 0.677 | 0.733 | 0.688 | 0.746 |
| $\mu$ | 0.737 | 0.737 | 0.694 | 0.779 | 0.829 | 0.846 | 0.820 | 0.874 | 0.668 | 0.691 | 0.630 | 0.717 |
| $\sigma$ | 0.017 | 0.019 | 0.027 | 0.014 | 0.043 | 0.041 | 0.023 | 0.036 | 0.009 | 0.028 | 0.043 | 0.023 |
| **Panel B: 10-Fold Cross-validation on AUC scores** | | | | | | | | | | | | |
| | Combined | | | | Corporate | | | | People | | | |
| **Round** | **LR** | **RF** | **SVM** | **GBM** | **LR** | **RF** | **SVM** | **GBM** | **LR** | **RF** | **SVM** | **GBM** |
| 1 | 0.769 | 0.757 | 0.721 | 0.799 | 0.870 | 0.902 | 0.865 | 0.915 | 0.655 | 0.692 | 0.571 | 0.708 |
| 2 | 0.777 | 0.770 | 0.760 | 0.814 | 0.798 | 0.811 | 0.780 | 0.864 | 0.664 | 0.697 | 0.570 | 0.718 |
| 3 | 0.719 | 0.722 | 0.674 | 0.761 | 0.823 | 0.818 | 0.804 | 0.870 | 0.651 | 0.677 | 0.595 | 0.702 |
| 4 | 0.746 | 0.754 | 0.720 | 0.812 | 0.820 | 0.824 | 0.823 | 0.844 | 0.689 | 0.672 | 0.622 | 0.685 |
| 5 | 0.710 | 0.693 | 0.692 | 0.762 | 0.810 | 0.819 | 0.810 | 0.867 | 0.670 | 0.657 | 0.570 | 0.746 |
| 6 | 0.754 | 0.726 | 0.714 | 0.763 | 0.870 | 0.867 | 0.818 | 0.907 | 0.686 | 0.761 | 0.642 | 0.774 |
| 7 | 0.743 | 0.736 | 0.696 | 0.769 | 0.866 | 0.888 | 0.847 | 0.917 | 0.691 | 0.735 | 0.683 | 0.766 |
| 8 | 0.726 | 0.740 | 0.713 | 0.786 | 0.807 | 0.850 | 0.819 | 0.869 | 0.653 | 0.651 | 0.550 | 0.669 |
| 9 | 0.724 | 0.736 | 0.738 | 0.769 | 0.847 | 0.865 | 0.804 | 0.877 | 0.648 | 0.660 | 0.572 | 0.670 |
| 10 | 0.729 | 0.728 | 0.702 | 0.775 | 0.807 | 0.854 | 0.788 | 0.894 | 0.700 | 0.738 | 0.663 | 0.750 |
| $\mu$ | 0.740 | 0.736 | 0.713 | 0.781 | 0.832 | 0.850 | 0.816 | 0.882 | 0.671 | 0.694 | 0.604 | 0.719 |
| $\sigma$ | 0.022 | 0.021 | 0.024 | 0.021 | 0.029 | 0.031 | 0.025 | 0.025 | 0.019 | 0.038 | 0.046 | 0.039 |

**Notes:** The table reports the AUCs for 5-fold and 10-fold cross-validation for the hybrid-sampled Country-level model with logistic regression (LR), random forest (RF), support vector machine (SVM) and gradient boosting (GBM). The data sample comprises of 82,964 alerts (30,303 corporate-related and 44,529 people-related) with 1,240 Issue cases (524 corporate-related and 659 people-related). The model has 12 predictors.

**Table 5: Country-level Predictor Importance for Country-level Model with Hybrid-sampling**

| | Combined | | | | Corporate | | | | People | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **LR** | **RF** | **GBM** | **Ave.** | **LR** | **RF** | **GBM** | **Ave.** | **LR** | **RF** | **GBM** | **Ave.** |
| **$CPI_R$** | * | 5 | 5 | 5 | *** | 5 | 4 | 4 | *** | 3 | 3 | 3 |
| **$FSI_R$** | *** | 6 | 3 | 4 | · | 8 | 6 | 8 | *** | 1 | 2 | 2 |
| **$IDV_R$** | *** | 1 | 1 | 1 | *** | 1 | 1 | 1 | *** | 2 | 4 | 4 |
| **$MAS_R$** | *** | 9 | 6 | 8 | | 9 | 11 | 9 | *** | 9 | 8 | 8 |
| **$PDI_R$** | *** | 4 | 7 | 6 | *** | 4 | 7 | 5 | *** | 6 | 6 | 5 |
| **$UAI_R$** | *** | 8 | 10 | 9 | *** | 7 | 5 | 6 | *** | 8 | 7 | 7 |
| **$CPI_W$** | *** | 3 | 4 | 3 | *** | 3 | 3 | 3 | *** | 4 | 1 | 1 |
| **$FSI_W$** | | 12 | 8 | 11 | *** | 10 | 12 | 12 | * | 12 | 11 | 12 |
| **$IDV_W$** | | 2 | 2 | 2 | *** | 2 | 2 | 2 | *** | 7 | 12 | 11 |
| **$MAS_W$** | *** | 10 | 11 | 10 | | 11 | 10 | 10 | | 5 | 10 | 9 |
| **$PDI_W$** | *** | 7 | 9 | 7 | * | 6 | 8 | 7 | *** | 11 | 9 | 10 |
| **$UAI_W$** | *** | 11 | 12 | 12 | *** | 12 | 9 | 11 | *** | 10 | 5 | 6 |

**Notes:** The table reports the importance of the Country-level predictors by ranking for the Hybrid-sampled Country-level model applied to the full sample (combined) and its partitions (Corporate & People accounts). Estimates of importance are obtained from the logistic regression (LR), random forest (RF), gradient boosted model (GBM) algorithms. A weighted average of RF and GBM (Ave.) is included. For LR, ***,**,* and · denote 0.1%, 1%, 5% and 10% levels of significance. RF and GBM are both tree-based algorithms and so their estimates are based on the mean decrease in the Gini index of each node across all trees. The Gini index measures node impurity. The data sample comprises of 82,964 alerts (30,303 corporate-related and 44,529 people-related) with 1,240 Issue cases (524 corporate-related and 659 people-related). The model has 12 predictors.

**Table 6: Country, Account & Transaction-level Models**

| | | Combined | | | Corporate | | | People | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Model** | **Balancing** | **TPR** | **FPR** | **AUC** | **TPR** | **FPR** | **AUC** | **TPR** | **FPR** | **AUC** |
| **LR** | **No Balancing** | 0.72 | 0.44 | 0.740 | 0.84 | 0.41 | 0.836 | 0.73 | 0.46 | 0.714 |
| | **Under-sampling** | 0.73 | 0.41 | 0.747 | 0.90 | 0.46 | 0.849 | 0.74 | 0.48 | 0.706 |
| | **Hybrid-sampling** | 0.72 | 0.42 | 0.747 | 0.90 | 0.54 | 0.846 | 0.69 | 0.43 | 0.712 |
| | **Synthetic-sampling** | 0.71 | 0.42 | 0.740 | 0.87 | 0.52 | 0.831 | 0.74 | 0.49 | 0.685 |
| **RF** | **No Balancing** | 0.96 | 0.53 | 0.908 | 0.92 | 0.28 | 0.930 | 1.00 | 1.00 | 0.842 |
| | **Under-sampling** | 0.93 | 0.48 | 0.895 | 0.97 | 0.55 | 0.932 | 0.91 | 0.59 | 0.835 |
| | **Hybrid-sampling** | 0.94 | 0.47 | 0.911 | 0.96 | 0.41 | 0.938 | 0.88 | 0.49 | 0.848 |
| | **Synthetic-sampling** | 1.00 | 1.00 | 0.772 | 0.89 | 0.42 | 0.877 | 1.00 | 1.00 | 0.638 |
| **SVM** | **No Balancing** | 0.88 | 0.51 | 0.835 | 0.91 | 0.59 | 0.881 | 0.73 | 0.48 | 0.737 |
| | **Under-sampling** | 0.89 | 0.51 | 0.801 | 0.95 | 0.54 | 0.880 | 0.79 | 0.57 | 0.739 |
| | **Hybrid-sampling** | 0.86 | 0.43 | 0.845 | 0.90 | 0.53 | 0.886 | 0.72 | 0.53 | 0.722 |
| | **Synthetic-sampling** | 0.65 | 0.41 | 0.723 | 0.86 | 0.52 | 0.847 | 0.55 | 0.40 | 0.603 |
| **GBM** | **No Balancing** | 0.88 | 0.47 | 0.842 | 0.93 | 0.59 | 0.880 | 0.84 | 0.57 | 0.769 |
| | **Under-sampling** | 0.93 | 0.53 | 0.853 | 0.95 | 0.40 | 0.916 | 0.82 | 0.40 | 0.799 |
| | **Hybrid-sampling** | 0.87 | 0.41 | 0.863 | 0.93 | 0.40 | 0.921 | 0.83 | 0.47 | 0.799 |
| | **Synthetic-sampling** | 0.67 | 0.40 | 0.717 | 0.88 | 0.40 | 0.846 | 0.76 | 0.55 | 0.652 |

**Notes:** The table reports the performance of our Country, Account & Transaction-level model using logistic regression (LR), random forest (RF), support vector machine (SVM) and gradient boosting (GBM) in combination with no balancing, under-sampling, hybrid-sampling and synthetic-sampling, respectively. The performance is measured using True Positive Rate (TP Rate), False Positive Rate (FP Rate) and Area under the ROC Curve (AUC). The data sample comprises of 74,246 alerts (30,292 corporate-related and 43,954 people-related) with 1,182 Issue cases (524 corporate-related and 648 people-related). The model has 24 predictors.

**Table 7: Country, Account & Transaction-level Models with PROP Score**

| | | Combined | | | Corporate | | | People | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Model** | **Balancing** | **TPR** | **FPR** | **AUC** | **TPR** | **FPR** | **AUC** | **TPR** | **FPR** | **AUC** |
| **LR** | **No Balancing** | 0.77 | 0.48 | 0.754 | 0.82 | 0.40 | 0.840 | 0.79 | 0.47 | 0.733 |
| | **Under-sampling** | 0.78 | 0.45 | 0.763 | 0.91 | 0.49 | 0.848 | 0.79 | 0.47 | 0.734 |
| | **Hybrid-sampling** | 0.77 | 0.44 | 0.764 | 0.90 | 0.58 | 0.851 | 0.76 | 0.46 | 0.733 |
| | **Synthetic-sampling** | 0.79 | 0.47 | 0.756 | 0.86 | 0.47 | 0.834 | 0.83 | 0.56 | 0.723 |
| **RF** | **No Balancing** | 0.89 | 0.40 | 0.894 | 0.95 | 0.33 | 0.946 | 1.00 | 1.00 | 0.845 |
| | **Under-sampling** | 0.87 | 0.43 | 0.878 | 0.95 | 0.40 | 0.943 | 0.91 | 0.57 | 0.846 |
| | **Hybrid-sampling** | 0.92 | 0.44 | 0.901 | 0.97 | 0.44 | 0.952 | 0.83 | 0.41 | 0.855 |
| | **Synthetic-sampling** | 0.90 | 0.58 | 0.790 | 0.88 | 0.43 | 0.873 | 1.00 | 1.00 | 0.690 |
| **SVM** | **No Balancing** | 0.87 | 0.50 | 0.828 | 0.89 | 0.54 | 0.883 | 0.78 | 0.59 | 0.742 |
| | **Under-sampling** | 0.88 | 0.57 | 0.789 | 0.94 | 0.51 | 0.896 | 0.81 | 0.57 | 0.753 |
| | **Hybrid-sampling** | 0.88 | 0.53 | 0.833 | 0.91 | 0.57 | 0.896 | 0.75 | 0.54 | 0.731 |
| | **Synthetic-sampling** | 0.78 | 0.56 | 0.745 | 0.83 | 0.41 | 0.848 | 0.65 | 0.49 | 0.667 |
| **GBM** | **No Balancing** | 0.89 | 0.56 | 0.829 | 0.92 | 0.58 | 0.886 | 0.91 | 0.57 | 0.818 |
| | **Under-sampling** | 0.87 | 0.43 | 0.850 | 0.94 | 0.40 | 0.922 | 0.91 | 0.51 | 0.828 |
| | **Hybrid-sampling** | 0.87 | 0.42 | 0.855 | 0.96 | 0.55 | 0.926 | 0.83 | 0.40 | 0.828 |
| | **Synthetic-sampling** | 0.74 | 0.48 | 0.709 | 0.88 | 0.48 | 0.840 | 0.80 | 0.58 | 0.689 |

**Notes:** The table reports the performance of our Country, Account & Transaction-level model, with the PROP score variable included, using logistic regression (LR), random forest (RF), support vector machine (SVM) and gradient boosting (GBM) in combination with no balancing, under-sampling, hybrid-sampling and synthetic-sampling, respectively. The performance is measured using True Positive Rate (TP Rate), False Positive Rate (FP Rate) and Area under the ROC Curve (AUC). The data sample comprises of 74,724 alerts (30,292 corporate-related and 43,954 people-related) with 1,182 Issue cases (524 corporate-related and 648 people-related). The model has 25 predictors.

**Table 8: Country-level Predictor Importance for Country, Account & Transaction-level Model with Hybrid-sampling**

| | Combined | | | | Corporate | | | | People | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **LR** | **RF** | **GBM** | **Ave.** | **LR** | **RF** | **GBM** | **Ave.** | **LR** | **RF** | **GBM** | **Ave.** |
| **$CPI_R$** | | 12 | 18 | 17 | *** | 12 | 18 | 12 | *** | 14 | 13 | 13 |
| **$FSI_R$** | *** | 8 | 6 | 7 | | 14 | 12 | 14 | *** | 9 | 9 | 9 |
| **$IDV_R$** | | 5 | 1 | 1 | *** | 2 | 1 | 1 | *** | 7 | 1 | 2 |
| **$MAS_R$** | *** | 15 | 11 | 13 | *** | 15 | 19 | 16 | *** | 12 | 17 | 16 |
| **$PDI_R$** | *** | 11 | 12 | 11 | *** | 9 | 10 | 10 | · | 16 | 19 | 17 |
| **$UAI_R$** | *** | 13 | 16 | 15 | *** | 5 | 4 | 4 | · | 15 | 14 | 15 |
| **$CPI_W$** | *** | 17 | 7 | 10 | *** | 10 | 14 | 11 | *** | 17 | 4 | 7 |
| **$FSI_W$** | *** | 21 | 17 | 20 | | 16 | 20 | 20 | ** | 23 | 22 | 24 |
| **$IDV_W$** | *** | 10 | 4 | 5 | | 4 | 2 | 3 | *** | 20 | 24 | 22 |
| **$MAS_W$** | * | 18 | 21 | 19 | | 18 | 17 | 19 | | 18 | 23 | 21 |
| **$PDI_W$** | *** | 22 | 20 | 21 | | 13 | 13 | 13 | *** | 24 | 21 | 23 |
| **$UAI_W$** | *** | 23 | 23 | 23 | *** | 21 | 23 | 21 | | 21 | 15 | 18 |

**Notes:** The table reports the importance of the Country-level predictors by absolute ranking for the Hybrid-sampled Country, Account & Transaction-level model applied to the full sample (combined) and its partitions (Corporate & People accounts). Estimates of importance are obtained from the logistic regression (LR), random forest (RF), gradient boosted model (GBM) algorithms. A weighted average of RF and GBM (Ave.) is included. For LR, ***,**,* and · denote 0.1%, 1%, 5% and 10% levels of significance. RF and GBM are both tree-based algorithms and so their estimates are based on the mean decrease in the Gini index of each node across all trees. The Gini index measures node impurity. The data sample comprises of 74,246 alerts (30,292 corporate-related and 43,954 people-related) with 1,182 Issue cases (524 corporate-related and 648 people-related). The model has 24 predictors.

**Table 9: Country-level Predictor Importance for Country, Account & Transaction-level Model with Hybrid-sampling and PROP Score included**

| | Combined | | | | Corporate | | | | People | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **LR** | **RF** | **GBM** | **Ave.** | **LR** | **RF** | **GBM** | **Ave.** | **LR** | **RF** | **GBM** | **Ave.** |
| **PROP** | *** | 5 | 3 | 3 | *** | 7 | 10 | 9 | *** | 4 | 1 | 1 |
| **$CPI_R$** | | 12 | 11 | 12 | *** | 14 | 12 | 13 | *** | 10 | 6 | 8 |
| **$FSI_R$** | *** | 9 | 6 | 8 | | 15 | 14 | 15 | *** | 8 | 4 | 7 |
| **$IDV_R$** | | 8 | 4 | 5 | *** | 3 | 2 | 2 | *** | 13 | 12 | 13 |
| **$MAS_R$** | *** | 19 | 17 | 17 | *** | 17 | 18 | 19 | *** | 15 | 13 | 16 |
| **$PDI_R$** | *** | 14 | 20 | 18 | *** | 11 | 8 | 10 | · | 14 | 16 | 15 |
| **$UAI_R$** | *** | 13 | 14 | 13 | *** | 12 | 11 | 11 | · | 17 | 17 | 17 |
| **$CPI_W$** | *** | 18 | 10 | 11 | *** | 13 | 6 | 6 | *** | 18 | 5 | 10 |
| **$FSI_W$** | *** | 24 | 19 | 21 | | 16 | 17 | 17 | ** | 24 | 20 | 24 |
| **$IDV_W$** | *** | 11 | 1 | 2 | | 2 | 1 | 1 | *** | 21 | 24 | 21 |
| **$MAS_W$** | * | 16 | 22 | 20 | | 18 | 16 | 16 | | 20 | 25 | 20 |
| **$PDI_W$** | *** | 21 | 21 | 22 | | 8 | 23 | 14 | *** | 25 | 19 | 25 |
| **$UAI_W$** | *** | 23 | 23 | 23 | *** | 22 | 20 | 22 | | 23 | 21 | 22 |

**Notes:** The table reports the importance of the Country-level predictors by absolute ranking for the Hybrid-sampled Country, Account & Transaction-level model applied to the full sample (combined) and its partitions (Corporate & People accounts) with the PROP score variable included. Estimates of importance are obtained from the logistic regression (LR), random forest (RF), gradient boosted model (GBM) algorithms. A weighted average of RF and GBM (Ave.) is included. For LR, ***,**,* and · denote 0.1%, 1%, 5% and 10% levels of significance. RF and GBM are both tree-based algorithms and so their estimates are based on the mean decrease in the Gini index of each node across all trees. The Gini index measures node impurity. The data sample comprises of 74,724 alerts (30,292 corporate-related and 43,954 people-related) with 1,182 Issue cases (524 corporate-related and 648 people-related). the model has 25 predictors.

## 3.10 Internet Appendices A-C

### 3.10.1 Internet Appendix A

**Table A1: Registration Type Profile**

| Reg Type | # Alerts | Alert Share | # Issues | Issue Share | Issue Rate |
|---|---|---|---|---|---|
| Corporate | 44,159 | 21.08 % | 936 | 38.13 % | 2.12 % |
| Education | 3,169 | 1.51 % | 10 | 0.41 % | 0.32 % |
| Estate-like | 670 | 0.32 % | 0 | 0.00 % | 0.00 % |
| IRA | 19,745 | 9.43 % | 14 | 0.57 % | 0.07 % |
| People | 119,717 | 57.15 % | 1,366 | 55.64 % | 1.14 % |
| Trust | 22,024 | 10.51 % | 129 | 5.25 % | 0.59 % |

**Notes:** The table reports the cross-section of Alerts and Issues over the different reg types that comprise the accounts which trigger the alerts. The categories of Corporate and People together compromise 78.23% of the alerts and 93.77% of the Issue cases in total and so, for the purposes of our study, we only consider these two reg types.

### 3.10.2 Internet Appendix B: Hofstede Indices

**Table B1: Hofstede Indices Models**

| | | Combined | | | Corporate | | | People | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Model** | **Balancing** | **TPR** | **FPR** | **AUC** | **TPR** | **FPR** | **AUC** | **TPR** | **FPR** | **AUC** |
| **LR** | **No Balancing** | 0.70 | 0.43 | 0.695 | 0.87 | 0.53 | 0.813 | 0.58 | 0.40 | 0.651 |
| | **Under-sampling** | 0.70 | 0.44 | 0.711 | 0.87 | 0.49 | 0.818 | 0.67 | 0.51 | 0.659 |
| | **Hybrid-sampling** | 0.70 | 0.42 | 0.711 | 0.87 | 0.50 | 0.818 | 0.65 | 0.48 | 0.663 |
| | **Synthetic-sampling** | 0.78 | 0.53 | 0.711 | 0.81 | 0.44 | 0.812 | 0.71 | 0.54 | 0.658 |
| **RF** | **No Balancing** | 1.00 | 1.00 | 0.573 | 1.00 | 1.00 | 0.664 | 1.00 | 1.00 | 0.514 |
| | **Under-sampling** | 0.71 | 0.42 | 0.747 | 0.86 | 0.42 | 0.848 | 0.75 | 0.49 | 0.718 |
| | **Hybrid-sampling** | 0.71 | 0.41 | 0.730 | 0.82 | 0.32 | 0.848 | 0.78 | 0.49 | 0.707 |
| | **Synthetic-sampling** | 1.00 | 1.00 | 0.713 | 1.00 | 1.00 | 0.835 | 1.00 | 1.00 | 0.654 |
| **SVM** | **No Balancing** | 0.67 | 0.53 | 0.545 | 0.51 | 0.45 | 0.532 | 0.72 | 0.53 | 0.625 |
| | **Under-sampling** | 0.74 | 0.56 | 0.670 | 0.92 | 0.52 | 0.830 | 0.62 | 0.41 | 0.648 |
| | **Hybrid-sampling** | 0.73 | 0.54 | 0.686 | 0.89 | 0.56 | 0.829 | 0.79 | 0.59 | 0.641 |
| | **Synthetic-sampling** | 0.60 | 0.59 | 0.531 | 0.62 | 0.41 | 0.719 | 0.66 | 0.58 | 0.586 |
| **GBM** | **No Balancing** | 0.82 | 0.52 | 0.765 | 0.89 | 0.57 | 0.867 | 0.82 | 0.56 | 0.727 |
| | **Under-sampling** | 0.81 | 0.52 | 0.767 | 0.84 | 0.41 | 0.861 | 0.82 | 0.56 | 0.726 |
| | **Hybrid-sampling** | 0.83 | 0.54 | 0.771 | 0.84 | 0.41 | 0.866 | 0.82 | 0.55 | 0.739 |
| | **Synthetic-sampling** | 0.68 | 0.41 | 0.716 | 0.85 | 0.57 | 0.811 | 0.71 | 0.51 | 0.662 |

**Notes:** The table reports the performance of our Hofstede Indices model with Cultural Distance using logistic regression (LR), random forest (RF), support vector machine (SVM) and gradient boosting (GBM) in combination with no balancing, under-sampling, hybrid-sampling and synthetic-sampling, respectively. The performance is measured using True Positive Rate (TP Rate), False Positive Rate (FP Rate) and Area under the ROC Curve (AUC). The data sample comprises of 81,858 alerts (32,482 corporate-related and 49,376 people-related) with 1,273 Issue cases (537 corporate-related and 736 people-related). The model has 8 predictors.

**Table B2: Predictor Importance for Hofstede Indices Model with Hybrid-sampling**

| | Combined | | | | Corporate | | | | People | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Predictor** | **LR** | **RF** | **GBM** | **Ave.** | **LR** | **RF** | **GBM** | **Ave.** | **LR** | **RF** | **GBM** | **Ave.** |
| **$IDV_S$** | *** | 1 | 1 | 1 | *** | 1 | 1 | 1 | *** | 1 | 2 | 1 |
| **$MAS_S$** | *** | 7 | 5 | 5 | | 7 | 6 | 6 | *** | 2 | 3 | 3 |
| **$PDI_S$** | . | 2 | 3 | 3 | *** | 2 | 3 | 3 | *** | 4 | 5 | 5 |
| **$UAI_S$** | *** | 4 | 4 | 4 | *** | 5 | 5 | 5 | *** | 5 | 4 | 4 |
| **$IDV_R$** | *** | 3 | 2 | 2 | | 3 | 4 | 4 | * | 6 | 8 | 7 |
| **$MAS_R$** | *** | 5 | 6 | 6 | *** | 6 | 8 | 8 | *** | 3 | 1 | 2 |
| **$PDI_R$** | . | 6 | 7 | 7 | *** | 4 | 2 | 2 | ** | 7 | 7 | 8 |
| **$UAI_R$** | * | 8 | 8 | 8 | *** | 8 | 7 | 7 | *** | 8 | 6 | 6 |

**Notes:** The table reports the importance of the predictors for the Hybrid-sampled Hofstede Indices model applied to the full sample (combined) and its partitions (Corporate & People accounts). Estimates of importance are obtained from the logistic regression (LR), random forest (RF), gradient boosted model (GBM) algorithms. A weighted average of RF and GBM (Ave.) is included. For LR, ***,**,* and · denote 0.1%, 1%, 5% and 10% levels of significance. RF and GBM are both tree-based algorithms and so their estimates are based on the mean decrease in the Gini index of each node across all trees. The Gini index measures node impurity. The data sample comprises of 81,858 alerts (32,482 corporate-related and 49,376 people-related) with 1,273 Issue cases (537 corporate-related and 736 people-related). The model has 8 predictors.

## Table B3: Cross-validation for Hofstede Indices Model with Hybrid-sampling

**Panel A: 5-Fold Cross-validation on AUC scores**

| | Combined | | | | Corporate | | | | People | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Round** | **LR** | **RF** | **SVM** | **GBM** | **LR** | **RF** | **SVM** | **GBM** | **LR** | **RF** | **SVM** | **GBM** |
| 1 | 0.701 | 0.745 | 0.736 | 0.772 | 0.855 | 0.877 | 0.848 | 0.873 | 0.647 | 0.691 | 0.629 | 0.717 |
| 2 | 0.734 | 0.761 | 0.721 | 0.796 | 0.814 | 0.859 | 0.794 | 0.876 | 0.648 | 0.664 | 0.618 | 0.714 |
| 3 | 0.710 | 0.741 | 0.728 | 0.778 | 0.805 | 0.837 | 0.801 | 0.863 | 0.644 | 0.684 | 0.653 | 0.685 |
| 4 | 0.717 | 0.744 | 0.695 | 0.773 | 0.787 | 0.842 | 0.821 | 0.884 | 0.640 | 0.731 | 0.680 | 0.756 |
| 5 | 0.692 | 0.731 | 0.700 | 0.768 | 0.774 | 0.825 | 0.780 | 0.842 | 0.694 | 0.703 | 0.680 | 0.721 |
| $\mu$ | 0.711 | 0.744 | 0.716 | 0.777 | 0.807 | 0.848 | 0.809 | 0.868 | 0.655 | 0.695 | 0.652 | 0.719 |
| $\sigma$ | 0.016 | 0.011 | 0.018 | 0.011 | 0.031 | 0.020 | 0.026 | 0.016 | 0.022 | 0.025 | 0.029 | 0.025 |

**Panel B: 10-Fold Cross-validation on AUC scores**

| | Combined | | | | Corporate | | | | People | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Round** | **LR** | **RF** | **SVM** | **GBM** | **LR** | **RF** | **SVM** | **GBM** | **LR** | **RF** | **SVM** | **GBM** |
| 1 | 0.757 | 0.764 | 0.689 | 0.785 | 0.760 | 0.810 | 0.751 | 0.851 | 0.658 | 0.668 | 0.618 | 0.699 |
| 2 | 0.745 | 0.771 | 0.714 | 0.822 | 0.795 | 0.836 | 0.799 | 0.875 | 0.716 | 0.690 | 0.634 | 0.704 |
| 3 | 0.706 | 0.735 | 0.703 | 0.766 | 0.772 | 0.789 | 0.781 | 0.825 | 0.618 | 0.691 | 0.612 | 0.710 |
| 4 | 0.682 | 0.745 | 0.729 | 0.774 | 0.773 | 0.826 | 0.795 | 0.884 | 0.664 | 0.735 | 0.650 | 0.740 |
| 5 | 0.762 | 0.775 | 0.736 | 0.817 | 0.706 | 0.770 | 0.756 | 0.821 | 0.654 | 0.687 | 0.626 | 0.720 |
| 6 | 0.690 | 0.723 | 0.701 | 0.733 | 0.891 | 0.903 | 0.867 | 0.931 | 0.653 | 0.734 | 0.655 | 0.747 |
| 7 | 0.729 | 0.764 | 0.716 | 0.779 | 0.865 | 0.907 | 0.869 | 0.898 | 0.665 | 0.777 | 0.636 | 0.767 |
| 8 | 0.697 | 0.728 | 0.694 | 0.772 | 0.828 | 0.889 | 0.872 | 0.896 | 0.644 | 0.670 | 0.639 | 0.688 |
| 9 | 0.703 | 0.757 | 0.720 | 0.781 | 0.814 | 0.864 | 0.844 | 0.898 | 0.640 | 0.735 | 0.667 | 0.723 |
| 10 | 0.648 | 0.684 | 0.655 | 0.744 | 0.850 | 0.903 | 0.840 | 0.900 | 0.616 | 0.683 | 0.638 | 0.703 |
| $\mu$ | 0.712 | 0.745 | 0.706 | 0.777 | 0.805 | 0.850 | 0.817 | 0.878 | 0.653 | 0.707 | 0.637 | 0.720 |
| $\sigma$ | 0.036 | 0.028 | 0.023 | 0.028 | 0.055 | 0.051 | 0.047 | 0.035 | 0.028 | 0.036 | 0.017 | 0.025 |

**Notes:** The table reports the AUCs for 5-fold and 10-fold cross-validation for the hybrid-sampled Hofstede Indices model with logistic regression (LR), random forest (RF), support vector machine (SVM) and gradient boosting (GBM). The data sample comprises of 81,858 alerts (32,482 corporate-related and 49,376 people-related) with 1,273 Issue cases (537 corporate-related and 736 people-related). The model has 8 predictors.

### 3.10.3 Internet Appendix C: Money Laundering

Pervasive across borders and undermining local economies, money laundering remains an issue of global concern. In generating and disbursing illicit proceeds from criminal activities that have integrated into the financial system, it paves the way for further financial illegal activity, compounding the problem. Money laundering is thus a channel to legitimise dirty money (i.e., money generated from illegal activities) by integrating it into any established financial system for subsequent use without evoking suspicion. In short, dirty money is transacted in a manner concealing or obscuring its criminal origins. Although difficult to measure with any degree of certainty, estimates for the total amount of money laundered worldwide range from 2-5% of global GDP (approximately $600 billion to $1.6 trillion).[69]

Combating money laundering requires cooperation between the public and private sectors. However, current compliance requirements, such as transaction monitoring and suspicious activity reporting, impose significant costs on the private sector with negligible returns.[70] Nowhere is the futility of this endeavour highlighted better than in the recent high-profile scandals involving Danske Bank and Swedbank. Further, the present AML surveillance is painstakingly inefficient, time-consuming, and labor-intensive. Financial institutions vet thousands of potentially suspicious transactions every day and any failure to comply with the anti-money laundering (AML) surveillance requirements often subjects them to substantial fines and penalties by the financial regulatory bodies. Financial institutions are continually increasing their investments to detect and curb money laundering. For instance, since 2000, the International Monetary Fund (IMF) has redoubled its work on AML. Following the tragic events of 11 September 2001, IMF has also expanded its activities to include combating the financing of terrorism (CFT). Further, it launched in 2009 a donor-supported trust fund to finance AML/CFT capacity development in its member countries.

In light of both the recent work on the role of culture in corporate misconduct and bank failure (Liu, 2016; Berger et al., 2019), we explore the relevance of several country-specific cultural and institution quality indices against modelling the incidence of suspicious money movement

[69]To estimate the amount of money laundered, inferences are drawn best from relevant data. An example of such data is the 2002 National Money Laundering Strategy, a report from 1999-2003 by the US Treasury on Anti-Money Laundering (AML) drive. According to this report, $386 million worth of assets were seized in relation to money laundering in 2001, with a corresponding figure of $241 million in forfeited assets. However, such sums are considered only a small fraction of the actual figure. Further, various techniques and schools of thought have been employed to estimate the extent of money laundering reliably and consistently. The macroeconomic approach holds that the demand for money laundering is related to the monetary component of the so-called shadow economy, and tools such as currency-demand analysis (Tanzi, 1980) prove useful in this regard. A study conducted by the United Nations Office on Drugs and Crime (UNODC) investigated the volume of illegal funds generated by drug trafficking and organized crime and to what extent these funds are laundered. Their findings estimated that in 2009, criminal activity amounted to 3.6% of global GDP with 2.7% being laundered, valued about $1.6 trillion.

[70]The Lexis Nexis Risk Solutions' 2018 report estimates that AML compliance costs US financial firms approximately $25 billion annually.

within a financial institution. Another country-specific measure incorporated in this study pertains to secrecy jurisdiction, a concept introduced by Cobham et al. (2015). They suggest that jurisdictions are situated across a spectrum of secrecy in terms of financial sector and global market share. Secrecy jurisdiction is in contrast to binary classification of Tax Haven/Offshore Finance. The concept of secrecy jurisdiction shifts the narrow tax focused narrative onto a broader sense of financial secrecy and transparency, which eventually may facilitate changes in policy and practice. This measure is particularly relevant in the context of money laundering activity detection, given that a jurisdiction with a higher level of secrecy in financial sector is more likely to attract higher volume of transactions initiated with the intent of concealing its illegal origin. Such a case would defy the regulations pertaining to its jurisdiction because of the difficulty in obtaining necessary information to trace the money trail.

Besides existing practices such as Know-Your-Customer (KYC), AML operations within the private sector could further benefit from incorporating geopolitical or regulatory information. Thus, the investigative resources could be concentrated on these money laundering hotspots. As noted in an IIF (Institute of International Finance) study, the potential benefits of applying machine learning in anti-money laundering operation bristle with several challenges. A few key aspects highlighted in the study include AML specific challenges such as data quality, obstacles regarding data sharing, and legacy/dated IT infrastructure; machine learning-specific challenges such as ML talents, generalisation of trained models and interpretation of results, among other issues. We have first-hand experience with some, if not all, of these challenges during different stages of this study, in particular on issues of data quality, legacy IT systems, data sharing and protection, and the problem of real-world data having extremely imbalanced classes. Nevertheless, we utilise the data to the best of our knowledge and obtain useful results. Our results provide insights and empirical evidence for financial institutions willing to benefit from incorporating machine learning and publicly available data to their existing data framework to enhance AML operation.

# Countering racial discrimination in algorithmic lending: A case for model-agnostic interpretation methods

**Abstract**

Evidence of the validity of Shapley model-agnostic explainable AI methods in real-world datasets is scant. In respect to racial discrimination in lending, we examine the usefulness of Global Shapley Value and Shapley-Lorenz methods to attain algorithmic justice. Using 157,269 loan applications from the Home Mortgage Disclosure Act data set in New York during 2017, we confirm that these methods, consistent with the parameters of a logistic regression model, reveal evidence of racial discrimination. Critically, we show that these explainable AI methods can enable a financial institution to select an opaque creditworthiness model which blends out-of-sample performance with ethical considerations.

**JEL Classification: C52, C55, C58**

**Keywords:** Machine learning, Model-agnostic global interpretation methods, Algorithmic injustice, Big-Data lending

## 4.1 Introduction

Socio-economic instability and ethical lapses attendant on algorithmic injustice are a growing concern in financial services, for regulators and for governments. US senators Warren and Jones, for instance, in a letter in 2017 have cautioned the Consumer Financial Protection Bureau, the Federal Deposit Insurance Corporation, the Federal Reserve Board, and the Office of the Comptroller of the Currency to the perils of algorithmic lending and the need to curb algorithmic injustice. While the advent of AI has meant faster, inexpensive and historically accurate lending decisions, its models often fail to enhance the decisions' accountability. As a result, regulators and various national agencies in the US (USACM, 2017) and Europe (European Commission, 2019) stress the value of algorithmic transparency and accountability.

We test if state-of-the-art model-agnostic explainable AI (XAI) methods, Global Shapley Value (GSV) and Shapley-Lorenz (SL) can uncover evidence of injustice in the bank lending space. Extant literature provides strong evidence of prejudicial lending decisions in the US. For instance, Munnell et al. (1996) study the difference in home loan rejection rates between Boston's Blacks and Whites. They observe that while the rejection rate for White applicants with property and personal characteristics similar to Blacks is 20%, for the latter it is 28%. Similarly, Blanchflower et al. (2003) investigate if racial discrimination is evident in the small-business credit market. They find that small businesses run by Blacks are nearly twice as likely to be denied credit as their White counterparts, even after accounting for differences in borrower characteristics. Courchane and Nickerson (1997) and Black et al. (2003) find, conditional on the loan interest rate, African American borrowers pay more in points than Whites. Cheng

et al. (2015) find that African American borrowers, on average, pay 29 basis points more than comparable White borrowers. Bartlett et al. (2021) note that the rejection rate for minority applicants for government-sponsored enterprises, GSEs, (Federal Housing Administration, FHA) conventional home purchase loans is 6.73% (9%) higher than for non-minority applicants; for refinance loans, minority applicants are denied credit 5.96% (6.8%) more than for non-minority loan applicants.

Since prior literature finds evidence of racial discrimination in both in-person and algorithmic lending against African Americans in the US (Black et al., 1978; Munnell et al., 1996; Blanchflower et al., 2003; Butler et al., 2020; Bartlett et al., 2022), we examine if Global Shapley Value and Shapley-Lorenz XAI methods can uncover algorithmic injustice in the bank lending space. Our study examines 157,269 loan applications in New York (NY) in 2017 from the Home Mortgage Disclosure Act (HMDA) dataset. We first deploy a transparent logistic regression (LR) model and examine if, in its parameters, there is evidence consistent with racial discrimination.[71] We then examine if the XAI methods give insight regarding racial discrimination, consistent with the LR model. Ultimately, our contribution is pragmatic in that it shows how financial institutions can select opaque and complex models (e.g. random forests and support vector machines) which are both accurate, accountable and ethically preferable specifications that can mitigate racial discrimination in credit-worthiness decisions.

Prior studies have tested the validity of Shapley value-based model-agnostic XAI methods on simulated datasets. For instance, Štrumbelj and Kononenko (2010); Štrumbelj and Kononenko (2014) test the usefulness of Shapley value-based feature importance measure in explaining the model's outcomes on several simulated datasets and find that their proposed approximation method yields efficient and accurate explanations. Similarly, Aas et al. (2021) report that their improved Kernel SHAP approximation method efficiently computes Shapley values and provides accurate explanations on simulated datasets. However, evidence for the methods' validity on real-world datasets is scant. Therefore, we extend the literature by examining the usefulness of Shapley value-based XAI methods, namely Global Shapley Value and Shapley-Lorenz, on real-world data. Thus, this study aims to provide initial evidence on the usefulness or otherwise of the said XAI techniques in respect to uncovering racial discrimination in the lending space.

We also note a paucity of studies applying model-agnostic XAI methods and, in particular, the Shapley Value methodology in the financial economics literature. Exceptional applications of the Shapley Value approach, include the measurement of a bank's interconnectedness as a key driver of a bank's systemic importance (Drehmann and Tarashev, 2013) and how an increase

[71] It is expected that this is so. Bartlett et al. (2022) note significant differences in the rejection rates for minority and non-minority loan applicants for government-sponsored enterprises' and Federal Housing Administration's conventional home purchase loans and refinance mortgages. Further, Butler et al. (2020) estimate that 80,000 minority loan applications are rejected every year due to racial discrimination. In the same vein, in small-business and home loan credit markets too, Munnell et al. (1996) and Blanchflower et al. (2003) report that Blacks are more likely to be denied credit as compared to Whites.

in bank size leads to a more than proportional increase in systemic importance (Tarashev et al., 2016).[72] To the best of our knowledge, our study is the first to examine the usefulness of GSV and SL XAI methods to render accountable modeling decisions, in the topical[73] and important bank lending space.

To compute the GSV XAI metric, we first compute the Shapley values, using the Štrumbelj and Kononenko (2014) approach, for each feature at each instance.[74] We aggregate the feature contributions, across all instances, to arrive at a feature importance measure influencing the model behavior. To compute the SL metric we follow Giudici and Raffinetti (2021). Different to the classical variance decomposition, the Lorenz Zonoid decomposition is robust to outlying observations in that it is based on explained mutual variability, i.e., on the mutual distance between all observations, rather than deviations from the mean. It, thus, decomposes predictive accuracy rather than individual predictions. The SL method is a normalised measure (in the Receiver Operating Characteristics framework) and is calculated at both the global and local levels, and can be considered as a natural extension of the standard Shapley approach.[75]

We provide a brief outline of our findings. In line with Bartlett et al. (2022) which shows racial discrimination in mortgage lending data, we find, using an LR model in data sets from several class balancing approaches, that a government approved loan application by a Black is between 43 and 59 percent more likely to be rejected by a financial institution compared to that of a White. Across our class sampling approaches, and solely using binary input features, an applicant's race is the second largest and statistically significant LR model coefficient. Critically, we find that the GSV and SL XAI techniques approximately corroborate this finding, regarding

[72]Colombo and Pelagatti (2020) do investigate the relative importance of variables in predicting movements in exchange rate models but use comparatively limited partial dependence plots and the permutation measure, relative to the Shapley Value approach.

[73]See, for example, the Finanical Times, February 13, 2022: UK regulators warn banks on use of AI in loan applications.

[74]Derived from coalitional game theory (Shapley, 1953a), the derivation of Shapley Values assumes, for each instance of a prediction, that each feature value is a "player" in a game with the prediction as the payout (Molnar, 2020). Theoretically, Shapley values are the average marginal contribution of a feature value across all possible coalitions of features. They represent, in this way, a 'fair' distribution of the credit related to the difference between the specific prediction and the average prediction. This renders them insightful XAI methods which can shed light on the models' internal logic. Additionally, owing to the Monotonicity property of Shapley Values, the local Shapley formulation can be easily extended to provide insights into the models' global behaviour. To compute the Global Shapley Value approach XAI method, we first compute the Shapley values, using Strumbelj and Kononenko's (2014) approximation method, for each feature at each instance. We then aggregate the feature contributions across all instances to arrive at a feature importance measure influencing the model behavior. To compute the Shapley Lorenz measure we adopt Giudici and Raffinetti (2021) methodology.

[75]Both Global Shapley Value and Shapley Lorenz are model-agnostic and global explainable AI methods. However, Shapley Lorenz variable importance method in relying on the Lorenz Zonoid decomposition is more generally applicable and combines predictive accuracy with explainability. Both the XAI methods are tools for determining the relative importance of variables that impact the phenomenon of interest. However, Shapley Lorenz method in relying on the Lorenz Zonoid decomposition, which is based on the mutual distance between all the observations rather than deviations from the mean, is more robust to outlying observations than the Global Shapley Value method. Therefore, the results of Shapley Lorenz method are more reliable than Global Shapley Value. Further, unlike Global Shapley Value, Shapley Lorenz is a normalized measure of variable importance. This further attest that Shapley Lorenz method is a better variable importance method than Global Shapley Value.

the marginal contribution of application race in mortgage lending decisions. They indicate that the importance of applicant race is, across class sampling approaches, consistently ranked in the top 2 or 3 of our 7 input explanatory variables.

Turning to the out-of-sample predictive performance of the LR model, in respect to our balanced samples of historical decisions to extend credit, it correctly predicts about 73 percent of the instances of mortgage loan extensions (TPR). As indicated above, race of applicant is of high importance for these LR predictions. Hence, we also fit Random Forest (RF), and Support Vector Machine (SVM) model specifications, across the class sampling approaches. Our results suggest that these approaches also give useful out-of-sample performance.

The RF and SVM models technically perform at least as well in respect to the Area under the Curve (AUC) performance metric, and markedly better regarding the False Positive Rate (FPR) performance metric, than does the LR model specification. AUC is the probability that a random positive example (someone who is offered mortgage credit) will be ranked above (have a higher propensity to be offered credit) a random negative example (someone who is declined mortgage credit). FPR is the proportion of persons who are not extended credit who were predicted by the model to receive credit. In respect to the True Positive Rate (TPR), the RF and SVM models have performances several percentage points below the LR model. For example, the RF model has a TPR 3 percentage points less than the LR model in the over-sampling balanced data-set. This difference grows to a maximum of 5 percentage points in the hybrid sampling balanced dataset. What is especially interesting is that, in the RF model, the GSV XAI method ranks the importance of applicant's race as $4^{th}$ of the 7 explanatory input variables while the SL method ranks applicant race as $6^{th}$ of the 7 explanatory input variables. As a result, the RF model can be deemed as providing comparable performance to the LR model but is, comparatively at least, ethically accountable. Critically, using the GSV and SL approaches we show that while the RF model gives accountable decisions with relatively low importance accorded to the applicant's race, the SVM models are ethically inferior in that they rely heavily on information relating to an applicant's race.

Our illustrative work suggests that these explainable AI methods can enable financial institutions to select an opaque creditworthiness model which blends out-of-sample performance with ethical considerations. Our study, in this way, demonstrates, for financial institutions, tools which they can use to avail of otherwise opaque machine learning models, and remain in line with recommendations from regulators (USACM, 2017; OSTP, 2016; European Commission, 2019; France, 2018; Villani, 2018).

The remainder of the paper is organized as follows. The next section presents the literature review. Section 3 presents the data set. Section 4 presents the econometric methodology. Section 5 presents our results. Section 6 concludes.

## 4.2 Literature Review

No dearth of literature exists on racial discrimination in the US credit market. Going back as early as Black et al. (1978) seminal study, the existing literature investigates the economic criteria which determine a lender's decision to extend mortgage loan to a potential borrower, and, more importantly, whether the applicant's race and gender at all inform such a decision. The authors note that the rejection rate for minority home loan applicants is higher than it is for White applicants, even when the former group's income is higher. Much similarly, Munnell et al. (1996) study the difference in rejection rates between African Americans and the Whites in Boston. They employ a more comprehensive dataset than Black et al. (1978) that in accounting for both the HMDA loan applications' data and additional borrower characteristics, such as credit history and LTV, qualifies as a reliable study. Munnell et al. (1996) observe that while the rejection rate for White applicants with property and personal characteristics similar to African Americans is 20%, for the latter it is 28%. Similarly, Blanchflower et al. (2003) investigate whether there is racial discrimination in small-business credit market. They find that small businesses run by African Americans are nearly twice as likely to be denied credit as their White counterparts, even after accounting for differences in borrower characteristics. Additionally, their varied robustness checks signal that the difference in the rejection rates is unlikely to be explained by the omitted variable bias.

More recently, Butler et al. (2020) have found that the loan approval rate for Black and Hispanic auto loan applicants is 1.5% lower than it is for Whites, even after controlling for applicant's credit-worthiness. Further, they estimate that 80,000 minority loans applications are rejected every year due to racial discrimination. Similarly, Bartlett et al. (2022) note that the rejection rate for minority applicants for GSE (FHA) conventional home purchase loans is 6.73% (9%) higher than for non-minority applicants; for refinance loans, minority applicants are denied credit 5.96% (6.8%) more than for non-minority loan applicants.

Further, the extant literature also finds evidence that minority borrowers pay significantly higher interest rates than non-minority borrowers ((Courchane and Nickerson, 1997), Black et al. (2003), Ghent et al. (2014), Cheng et al. (2015), Zhang and Willen (2021), Bartlett et al. (2021)).

In this paper, we discuss two techniques for detecting discrimination in algorithmic lending. The socio-economic instability attendant on algorithmic injustice is a growing concern for regulators and governments. This study appropriately employs state-of-the-art model-agnostic explainable AI methods, namely Global Shapley Value and Shapley-Lorenz methods which are amenable to uncovering algorithmic injustice in the bank lending space. Thus, we promote the use of Global Shapley Value and Shapley-Lorenz explainable methods as a first check to detect discrimination in the models' outcomes. Specifically, we use the explainable methods to investigate whether an applicant's race determines the decision of the lender to accept/reject her

loan request. We acknowledge that our study may be plagued with omitted variable bias since we employ only the data on loan applications provided by HMDA without augmenting it with relevant borrower characteristics. However, the purpose of this study consists in demonstrating the usefulness of explainable model-agnostic methods in algorithmic lending.

## 4.3 Data and Variables

We examine 157,269 loan applications from HMDA's website made in NY during 2017. The dependent variable, *Declined Loan*, takes the value 1 if a loan application initially satisfies the approval requirements of GSEs/FHA, though it subsequently fails in meeting the lenders' requirements; it takes the value 0 if the lender approves the loan. Since the Global Financial Crisis, many lenders enforce stringent approval requirements besides those of GSEs and FHA. This means despite satisfying the requirements of GSEs/FHA, an applicant's loan request may still be rejected.[76] In detailing GSEs/FHA's initial acceptance of the borrower's application and its subsequent rejection by the bank, HMDA dataset that includes information on the applicant's race eminently qualifies for our study. Our key independent variable of interest is the information on applicant's race and we control for applicant's gender, income, amount of loan, purpose of loan, lien status, and type of loan. Concise definitions are provided in Table 1.

[Please insert Table 1 about here.]

Descriptive statistics and the correlation between the dependent and independent variables are reported in Table 2.

[Please insert Table 2 about here.]

We note that the dependent variable suffers from class imbalance. In other words, the number of observations that belong to the positive class (loan declined) is significantly lesser than those that belong to the negative class (loan approved). Models trained on such data in prioritizing the prevalent class over the minority class leads to an overly optimistic measure of accuracy (Batista et al., 2004). While such models can predict loan approvals with high level of accuracy, they often fail to accurately predict declined loans. Since the dependent variable suffers from severe class imbalance, we employ over-, under-, and hybrid-sampling techniques to meaningfully infer information from the data. Below, we discuss the resampling techniques employed in our study.

1. **Over sampling:** This technique randomly duplicates observations from the minority class to match the majority class size. This technique can be computationally expensive

[76]There are two stages in the loan process in the US. In the first stage, the lender submits the loan applicant's data (credit score, liquidity, debt-to-income ratio, LTV, property value etc.) to FHA or one of the two GSEs' automated underwriter systems (Desktop Underwriter for Fannie Mae; Loan Prospector for Freddie Mac). In the second stage, if the underwriter system issues an approval on the application, the lender can then decide whether or not to make an offer. Discrimination, therefore, occurs at the second stage of the loan application process.

(in cases of severe class imbalance, it may almost double the size of the dataset) and may lead to overfitting the model.

2. **Under sampling:** This technique randomly discards observations from the majority class to better balance the skewed distribution. In reducing the majority class's size to match the minority class, this technique, however, forgoes potentially useful information from the majority class.

3. **Hybrid sampling:** Combining under-sampling and over-sampling methods, this technique applies under-sampling technique to the majority class and over-sampling technique to the minority class to balance the class distribution.

## 4.4 Econometric methodology

In this section, we describe the model-agnostic variable importance measures which can be employed to uncover algorithmic injustice in the bank lending space.

### 4.4.1 Global Shapley Value Variable Importance Measure

Drawing on the fundamental coalitional Game Theory concepts, Shapley value variable importance measure quantifies the contribution of a feature in predicting the response value for a given instance. We aggregate the contributions of a feature across all instances to arrive at a measure that is interpreted as a measure of feature importance influencing the model behavior. In doing so, we treat our aggregated Shapley value of a feature, referred to as Global Shapley value, as a global model-agnostic variable importance measure. Here we introduce coalitional Game Theory concepts and then discuss the Shapley value model-agnostic measure.

A coalitional game is defined as a tuple $< N, \nu >$, where $N = \{1, 2, \ldots, n\}$ is a finite set of players and $\nu : 2^N \rightarrow \Re$ a characteristic function such that $\nu(\phi) = 0$. In the given definition, $N$ is referred to as the "grand coalition" of all the $n$ players and its subsets as coalitions, respectively. Defining each coalition's worth by the characteristic function, $\nu$, we seek to divide the value of the grand coalition, $\nu(N)$, in a "fair" manner among the individual players, assuming the grand coalition forms.

For a coalitional game that at least has a single player, there exists infinitely many solutions, such that some solutions are "fairer" than others. Here, a solution is an operator, $\Phi$, that assigns the tuple, $< N, \nu >$, a vector of payoffs, $\Phi(\nu) = (\Phi_1, \ldots, \Phi_n) \in \Re^n$. To axiomatize the notion of "fairness" of a solution, a "fair" solution must satisfy the following statements,
**Axiom 1** (Efficiency axiom): $\sum_{i \in N} \Phi_i(\nu) = \nu(N)$
**Axiom 2** (Symmetry axiom): For every coalition $S$ ($S \subset N$), if $\nu(S \cup \{i\}) = \nu(S \cup \{j\})$ holds

true for some players $i$ and $j$ such that $i, j \notin S$, then, $\Phi_i(v) = \Phi_j(v)$.
**Axiom 3** (Dummy axiom): For every coalition $S$ ($S \subset N$), if $v(S \cup \{i\}) = v(S)$ holds true, where $i \notin S$, then, $\Phi_i(v)$=0.
**Axiom 4** (Additivity axiom): For any pair of coalitional games $v$ and $\omega$, $\Phi(v+\omega) = \Phi(v) + \Phi(\omega)$, where $(v+\omega) = v(S) + \omega(S)$ for all coalitions $S$.

Shapley (1953b) proved that a unique solution exists for the coalitional game $< N, v >$, which satisfies all the four axioms and designated this unique solution as the Shapley value,

$$Shapley\ value_i(v) = \sum_{S \subseteq N \setminus \{i\}, s=|S|} \frac{(n-s-1)!\,s!}{n!}(v(S \cup \{i\}) - v(S)) \tag{44}$$

Štrumbelj and Kononenko (2014) developed the idea of Shapley value as a feature importance measure which I discuss below.

Let $A$ represent the feature space, $N = \{1, 2, \ldots, n\}$ be the $n$ features, and $f$ the classifier. Further, let $c$ be the class label with respect to which we intend explaining the prediction for the instance $x = (x_1, x_2, \ldots, x_n) \in A$. To quantify the contribution of a feature in the prediction of an instance, Štrumbelj and Kononenko (2014) developed the notion of a feature's contribution to the prediction difference between the classifier's prediction for the instance and expected prediction when the feature values are ignored. If $S$ is an arbitrary subset of $N$ ($S \subseteq N$), then this prediction difference can be generalized to $S$ as follows,

$$\Delta(S) = \frac{1}{|A_{N \setminus S}|} \sum_{y \in A_{N \setminus S}} f_c(\tau(x, y, S)) - \frac{1}{|A_N|} \sum_{y \in A_N} f_c(y) \tag{45}$$

$$\tau(x, y, S) = (z_1, z_2, ..., z_n);\ z_i = \begin{cases} x_i; & \text{if } i \in S \\ y_i & \text{if } i \notin S \end{cases}$$

In this generalization, only the features in $S$ are known and $\Delta(S)$ is the prediction difference between the expected prediction when the features belong to $S$ and the expected prediction when feature values do not belong to $S$. Štrumbelj and Kononenko (2014) also account for the interaction effects by implicitly defining each prediction difference $\Delta(S)$ to comprise $2^N$ contributions of interactions $I$,

$$\Delta(S) = \sum_{W \subseteq S} I(W),\ \ S \subseteq N \tag{46}$$

This is a recursive definition if we assume that the interaction of a null set is always zero,

$I(\phi) = 0$. This ensures that the interactions exist, and those can be uniquely defined as follows,

$$I(S) = \Delta(S) - \sum_{W \subset S} I(W), \;\; S \subseteq N \tag{47}$$

Given that each feature in an interaction shares an equal weight, we can distribute the interaction contributions among the n features as follows,

$$\varphi_i(\Delta) = \sum_{W \subseteq N \setminus \{i\}} \frac{I(W \cup \{i\})}{|W \cup \{i\}|}, \;\; i = 1, 2, ..., n \tag{48}$$

Since the Shapley values satisfy Axiom 1 (Shapley, 1953b), these feature contributions are implicitly normalized. This ensures both meaningful interpretation of the contributions and their easy comparison with one another. The said formulation, however, is computationally infeasible. Therefore, the authors suggest an efficient approximation, which we follow in our paper.

If we suppose $\Pi(N)$ to be the set of all ordered permutations of features, $N$, and define $Pre^i(O)$ as the set of features that precede feature, say, $i$, in the order $O \in \Pi(N)$. Then Castro et al. (2009) suggest an alternative formulation of the feature contributions as follows,

$$\varphi_i(\Delta) = \frac{1}{n!} \sum_{O \in \Pi(N)} (\Delta(Pre^i(O) \cup \{i\}) - \Delta(Pre^i(O))), \;\; i = 1, 2, ..., n \tag{49}$$

However, this formulation too is computationally infeasible when applied in the context of Štrumbelj and Kononenko (2014). To sidestep this problem, they extend the sampling algorithm to arrive at a prediction difference formulation that is equivalent to definition (45),

$$\Delta(S) = \frac{1}{|A|} \sum_{y \in A} (f(\tau(x, y, S)) - f(y)) \tag{50}$$

and substitute this definition in place of $\Delta$ in the formulation of Castro et al. (2009),

$$\varphi_i(\Delta) = \frac{1}{n!\,|A|} \sum_{O \in \Pi(N)} \sum_{y \in A} (f(\tau(x, y, Pre^i(O) \cup \{i\})) - f(\tau(x, y, Pre^i(O)))) \tag{51}$$

In this sampling procedure, $\Pi(N) \times A$ is the sampling population such that each order/instance pair defines a sample,

$$X_{O, y \in A} = f(\tau(x, y, Pre^i(O) \cup \{i\})) - f(\tau(x, y, Pre^i(O)))) \tag{52}$$

If randomly drawn, then all the samples have an equal chance, $(\frac{1}{n!|A|})$ , of being drawn, and this results in $E[X_{O, y \in A}] = \varphi_i$.

If we draw $m$ samples, with replacement, then $\hat{\varphi}_i = \frac{1}{m}\sum_{j=1}^{m} X_j$, where $X_j$ is the $j$th sample. Central Limit Theorem ensures that $\hat{\varphi}_i$ is normally distributed with mean $\varphi_i$ and variance $\frac{\sigma_i^2}{m}$, where $\sigma_i^2$ is the $i$th feature's population variance. Hence, we get a consistent and unbiased estimator, $\hat{\varphi}_i$, of $\varphi_i$.

Once we compute the Shapley values of a feature, say, $i$, across all the instances, we aggregate these contributions to arrive at a Global Shapley variable importance measure of the feature, $i$.

### 4.4.2 Shapley Lorenz Decomposition Variable Importance Measure

Developed by Giudici and Raffinetti (2021), the Shapley-Lorenz variable importance measure utilizes Lorenz Zonoid decompositions in the local Shapley value formulation to create a global model-agnostic explainable AI metric. Before we discuss the method, it would be in order to understand the Lorenz Zonoid decomposition and the Partial Gini Contribution measure (Giudici and Raffinetti, 2020).

Introduced by Koshevoy and Mosler (1996), Lorenz Zonoid is a generalization of Lorenz curve in higher dimensions, d. Specifically, in a multi-dimensional setting, Lorenz Zonoid is a generalization of the ROC curve; in a one-dimensional setting, it shows correspondence with the Gini coefficient.

Consider a training dataset $(X_i, Y_i)_{i=1}^{n}$, where $X = \{X_1, X_2, \ldots, X_K\}$ is the set of features and Y the response variable and $\hat{f}(X)$, the trained model. Then, the Lorenz Zonoid of Y and $\hat{f}(X)$ can be written as follows,

$$LZ_{d=1}(Y) = \frac{2Cov(Y, r(Y))}{n\mu} \tag{53}$$

and

$$LZ_{d=1}(\hat{f}(X))) = \frac{2Cov(\hat{f}(X), r(\hat{f}(X)))}{n\mu} \tag{54}$$

where  is the mean of the response variable Y, and r(Y) and r($\hat{f}(X)$) denote the rank scores of Y and f (X), respectively. Utilizing Lorenz Zonoid decompositions,Giudici and Raffinetti (2020) develop a dependence measure, Partial Gini Contribution (PGC), that quantifies the additional contribution of a feature to the existing model. More concretely, they define the PGC measure as follows,

$$PGC_{Y,X_h|X\setminus X_h} = \frac{LZ_{d=1}(\hat{f}(X))) - LZ_{d=1}(\hat{f}(X \setminus X_h))}{(LZ_{d=1}(Y) - LZ_{d=1}(\hat{f}(X \setminus X_h)} \tag{55}$$

Applying the numerator of the PGC measure as a pay-off function in the local Shapley value formulation, Giudici and Raffinetti (2021) arrive at the following global model-agnostic Shapley-Lorenz variable importance measure,

$$LZ_{d=1}^{X_k}(\hat{Y}) = \sum_{X' \subseteq C(X)\setminus X_k} \frac{|X'|!\,(K-|X'|-1)!}{K!}[LZ_{d=1}\hat{f}(X' \cup X_k) - LZ_{d=1}\hat{f}(X')] \tag{56}$$

In the given equations, the marginal contribution of $X_k, LZ_{d=1}^{X_k}(\hat{Y})$ is computed by considering all the possible model configurations excluding the variable $X_k$.

Although a global variable importance measure, the Shapley-Lorenz value in being easily applied to subsets of the total observations renders its applicability as a local variable importance measure.

## 4.5 Empirical findings

Table 3 reports the LR model's coefficient estimates on data balanced by the three data-balancing methods.[77] Across our class sampling approaches, applicant's race is the second largest and statistically significant LR model coefficient. Our results confirm that a Black's loan application is far more likely to be rejected than a White's. We find that a government approved loan application by a Black is between 43 and 59 percent more likely to be rejected by a financial institution compared to that of a White.[78]

[Please insert Table 3 about here.]

To ascertain if the XAI methods give insights consistent with the LR model, these methods are applied on the trained LR model, trained on 70% of the data. The results are reported in Table 4. Critically, we find that the GSV and SL XAI techniques corroborate the LR models' results, regarding the marginal contribution of application race in mortgage lending decisions. They indicate that the importance of applicant race is, across class sampling approaches, always ranked in the top 2 or 3 of our 7 explanatory variables. For the GSV method on the over-sampled dataset, *Loan Purpose* influences the model's decision optimally followed by *Applicant race*. Following *Applicant race*, the features that influence the model's decision, in decreasing order, are: *Loan type*, *Loan amount*, *Lien status*, *Applicant gender*, and *Applicant income*. Similarly, the SL feature importance measure determines *Loan Purpose* as the most

[77] Please see Table A1 in the Internet Appendix A for the marginal coefficient results of the LR models.

[78] (exp(0.4648)-1)*100; (exp(0.3600)-1)*100; (exp(0.4581)-1)*100

important feature followed by *Loan type*, *Applicant race*, *Applicant income*, *Lien status*, *Applicant gender*, and *Loan amount*. We note similar results for LR model on under-sampled and hybrid data. Consistent with the LR model results, the XAI methods reveal that an applicant's race determines the outcome of her loan application much more than other relevant criteria such as the loan amount and applicant's income. This testifies to the efficacy of XAI methods in uncovering discrimination.

[Please insert Table 4 about here.]

Further, we evaluate the out-of-sample predictive performance of the LR model, in respect to our sample of historical decisions to extend credit. We employ true positive rate (TPR), false positive rate (FPR), and AUC (area under the ROC curve) to evaluate the performance of the models. TPR measures the proportion of loan extensions correctly classified by the model and FPR measures the proportion of rejected loans misclassified by the model. To measure the model's out-of-sample predictive performance we compute the area under the ROC curve (AUC). AUC lies between 0 and 1. A model with AUC of 0.5 is no better than randomly guessing (random classifier) the class for an observation; a model with AUC less than 0.5 performs worse than the random classifier; and a model with AUC greater than 0.5 demonstrates predictive capacity. We find that the LR model correctly predicts about 73 percent of the instances of mortgage loan extensions (TPR). As indicated above, applicant's race is of high importance for these LR predictions. Hence, we also fit Random Forests, and Support Vector Machines, across the class sampling approaches. These models generally perform at least as well in respect to the AUC performance metrics, and markedly better regarding FPR. In respect to TPR, the RF and SVM models have performances several percentage points below the LR model. Table 5 reports these results.

[Please insert Table 5 about here.]

Further, the XAI methods applied on a RF model yields ethically accountable decisions. In the RF model, the GSV XAI method ranks the importance of applicant's race as 4/7 while the SL method ranks applicant race as 6/7. As a result, the RF model can be deemed as providing comparable performance to the LR model but is, comparatively at least, ethically accountable. In SVM model, however, we find that the *Applicant race* determines the decision to accept/reject an application. Our exploratory work suggests that these explainable AI methods can enable financial institutions to select an opaque creditworthiness model which blends out-of-sample performance with ethical considerations. Our study, in this way, illustrates, for financial institutions, tools which they can use to avail of otherwise opaque machine learning models, and remain in line with recommendations from regulators (USACM, 2017; OSTP, 2016; European Commission, 2019; France, 2018; Villani, 2018).

[Please insert Table 6 about here.]

## 4.6 Conclusion

In this study, we show that a transparent regression model specification provides evidence consistent with racial discrimination in mortgage lending decisions in New York. In this setting of racial discrimination, we then show that the Global Shapley Value and Shapley-Lorenz explanatory AI techniques give insights consistent with this principal finding of our transparent model, demonstrating their validity in a topical real world data set.

Of pragmatic importance, the study then illustrates the usefulness of these explanatory AI techniques. They are shown to uncover racial discrimination across opaque and sophisticated predictive models, and accordingly they permit a model specification selection which can obtain out-of-sample performance and which is ethically accountable.

## 4.7 Tables

**Table 1: Variable Definitions**

| Variable | Definition |
|---|---|
| Dependent variable | |
| Declined Loan | The variable takes the value 1 if a loan application initially satisfies the approval requirements of Government-sponsored enterprise (GSEs) or Federal Housing Administration (FHA), but it fails in meeting the lenders' requirements. It takes the value 0 if the lender approves the loan. |
| Independent variable of interest | |
| Applicant race | The variable takes the value 1 if the applicant is African American and 0 if White. |
| Control variables | |
| Applicant gender | The variable takes the value 1 if the applicant is male and 0 if female. |
| Applicant income | The variable takes the value 1 if the applicant's gross annual income is less than the median value and 0, otherwise. |
| Loan amount | The variable takes the value 1 if the loan amount is less than the median value and 0, otherwise. |
| Loan purpose | The variable takes the value 1 if the purpose of seeking a loan was for refinancing the mortgage and 0 if purchasing a home. |
| Lien status | This variable takes the value 1 if the loan application is secured by first lien and 0 for a sub-ordinate lien. A first lien is the first to be paid when a borrower defaults and the property or asset is used as collateral for the debt. |
| Loan type | The variable takes the value 1 if the loan was insured by the FHA and 0 if a conventional loan type. |

**Table 2: Descriptive Statistics**

| Variable | N | Min | Mean | Max | Std.Dev |
|---|---|---|---|---|---|
| Dependent variable | | | | | |
| Declined Loan | 157269 | 0 | 0.06 | 1 | 0.24 |
| Independent variable of interest | | | | | |
| Applicant race | 157269 | 0 | 0.08 | 1 | 0.27 |
| Control variables | | | | | |
| Applicant gender | 157269 | 0 | 0.65 | 1 | 0.48 |
| Applicant income | 157269 | 0 | 0.50 | 1 | 0.50 |
| Loan amount | 157269 | 0 | 0.87 | 1 | 0.34 |
| Loan purpose | 157269 | 0 | 0.33 | 1 | 0.47 |
| Lien status | 157269 | 0 | 0.97 | 1 | 0.16 |
| Loan type | 157269 | 0 | 0.18 | 1 | 0.38 |

**Panel A: Summary Statistics**

| | Applicant race | Applicant gender | Applicant income | Loan amount | Loan purpose | Lien status | Loan type |
|---|---|---|---|---|---|---|---|
| Correlation | 0.0440 | 0.0038 | 0.0099 | 0.0076 | 0.1200 | 0.0120 | 0.0330 |
| Standard Error | 0.0025 | 0.0025 | 0.0025 | 0.0025 | 0.0025 | 0.0025 | 0.0025 |
| t-statistic | 17.4299 | 1.5051 | 3.9271 | 3.0002 | 46.5735 | 4.6294 | 13.2530 |
| p-value | 0.0000 | 0.1320 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

**Panel B: Correlation Matrix**

Notes. Panel A presents summary statistics of the variables and Panel B presents the correlation coefficients of independent variables with the dependent variable. The associated standard errors, t-statistics, and p-values are also reported in Panel B. The variables are defined in Table 1.

**Table 3: Coefficient estimates of the Logistic Regression model**

| Variable | Estimate | Std. Error |
|---|---|---|
| Intercept | -0.5102*** | 0.0153 |
| Applicant race | 0.4648*** | 0.0159 |
| Applicant gender | 0.0681*** | 0.0097 |
| Applicant income | 0.0909*** | 0.0100 |
| Loan amount | -0.1007*** | 0.0150 |
| Loan purpose | 0.9659*** | 0.0093 |
| Lien status | -0.5970*** | 0.0307 |
| Loan type | 0.3094*** | 0.0118 |

**Panel A: Coefficient estimates on over-sampled dataset**

| Variable | Estimate | Std. Error |
|---|---|---|
| Intercept | -0.4600*** | 0.0586 |
| Applicant race | 0.3600*** | 0.0601 |
| Applicant gender | 0.0928* | 0.0372 |
| Applicant income | 0.1089** | 0.0384 |
| Loan amount | -0.1735** | 0.0577 |
| Loan purpose | 0.9775*** | 0.0356 |
| Lien status | -0.5753*** | 0.1217 |
| Loan type | 0.2757*** | 0.0455 |

**Panel B: Coefficient estimates on under-sampled dataset**

| Variable | Estimate | Std. Error |
|---|---|---|
| Intercept | -0.5077*** | 0.0210 |
| Applicant race | 0.4581*** | 0.0219 |
| Applicant gender | 0.0728*** | 0.0133 |
| Applicant income | 0.0975*** | 0.0137 |
| Loan amount | -0.1094*** | 0.0206 |
| Loan purpose | 0.9600*** | 0.0127 |
| Lien status | -0.5672*** | 0.0423 |
| Loan type | 0.3048*** | 0.0162 |

**Panel C: Coefficient estimates on hybrid-sampled dataset**

Notes. The Table presents the results of Logistic Regression Model. Our dependent variable is *Declined Loan*. All the variables are defined in Table 1 and we use the following significance stars * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

**Table 4: Feature Importance**

| Variable | Global Shapley | Rank | Shapley-Lorenz | Rank |
|---|---|---|---|---|
| Applicant race | 9538.816 | 2 | 18.654 | 3 |
| Applicant gender | -1998.523 | 6 | 3.997 | 6 |
| Applicant income | 204.311 | 7 | 5.299 | 4 |
| Loan amount | -5693.699 | 4 | 1.337 | 7 |
| Loan purpose | -53261.329 | 1 | 134.100 | 1 |
| Lien status | -2844.029 | 5 | 4.445 | 5 |
| Loan type | 6368.207 | 3 | 19.660 | 2 |

**Panel A: Marginal contribution of each explanatory variable in the Logistic Regression model on over-sampled dataset**

| Variable | Global Shapley | Rank | Shapley-Lorenz | Rank |
|---|---|---|---|---|
| Applicant race | 505.116 | 1 | 11.688 | 3 |
| Applicant gender | 67.133 | 5 | 2.831 | 5 |
| Applicant income | 16.387 | 6 | 5.061 | 4 |
| Loan amount | 304.712 | 3 | 1.057 | 7 |
| Loan purpose | 466.678 | 2 | 94.497 | 1 |
| Lien status | -173.152 | 4 | 2.829 | 6 |
| Loan type | 0.331 | 7 | 13.685 | 2 |

**Panel B: Marginal contribution of each explanatory variable in the Logistic Regression model on under-sampled dataset**

| Variable | Global Shapley | Rank | Shapley-Lorenz | Rank |
|---|---|---|---|---|
| Applicant race | 4904.751 | 2 | 9.636 | 3 |
| Applicant gender | -1136.422 | 5 | 2.312 | 5 |
| Applicant income | -3084.684 | 3 | 4.082 | 4 |
| Loan amount | 308.684 | 6 | 0.864 | 7 |
| Loan purpose | -17690.703 | 1 | 76.123 | 1 |
| Lien status | -1412.316 | 4 | 2.223 | 6 |
| Loan type | -34.382 | 7 | 10.927 | 2 |

**Panel C: Marginal contribution of each explanatory variable in the Logistic Regression model on hybrid-sampled dataset**

Notes. The Table presents the marginal contribution of each explanatory variable in the Logistic Regression model in terms of the standard Shapley approach, $\phi(\hat{f}(X_i))$, and linear Shapley-Lorenz approach, $LZ_{d=1}^{X_k}(\hat{y})$, respectively.

**Table 5: Out-of-sample predictive performance of Models on data balanced by various balancing methods**

| Model | TPR | FPR | AUC |
|---|---|---|---|
| LR | 73% | 50% | 64% |
| RF | 70% | 36% | 67% |
| SVM | 68% | 37% | 65% |

**Panel A: Data balanced by over-sampling method**

| Model | TPR | FPR | AUC |
|---|---|---|---|
| LR | 72% | 50% | 63% |
| RF | 68% | 43% | 63% |
| SVM | 65% | 40% | 62% |

**Panel B: Data balanced by under-sampling method**

| Model | TPR | FPR | AUC |
|---|---|---|---|
| LR | 74% | 52% | 63% |
| RF | 69% | 37% | 66% |
| SVM | 69% | 38% | 65% |

**Panel C: Data balanced by hybrid-sampling method**

Notes. The Table reports the performance of the Logistic Regression (LR), Random Forests (RF), and Support Vector Machine (SVM) models on over-, under-, and hybrid-sampled data. Over-, under-, and hybrid-sampling methods create balanced samples by randomly over-sampling minority examples, under-sampling majority examples and by combining over- and under-sampling, respectively. The performance is measured using True Positive Rate (TPR), False Positive Rate (FPR) and Area under the ROC Curve (AUC).

**Table 6: Feature Importance**

| Variable | Global Shapley | Rank | Shapley-Lorenz | Rank |
|---|---|---|---|---|
| Applicant race | 238.6 | 4 | 0.011 | 6 |
| Applicant gender | 193.5 | 5 | 0.029 | 5 |
| Applicant income | -49.3 | 7 | 0.163 | 2 |
| Loan amount | 2357.1 | 1 | 0.004 | 7 |
| Loan purpose | 563.4 | 2 | 0.097 | 3 |
| Lien status | -54.4 | 6 | 0.164 | 1 |
| Loan type | 318.1 | 3 | 0.029 | 4 |

**Panel A: Marginal contribution of each explanatory variable in the RF model on under-sampled dataset**

| Variable | Global Shapley | Rank | Shapley-Lorenz | Rank |
|---|---|---|---|---|
| Applicant race | 771.1 | 2 | 0.068 | 3 |
| Applicant gender | 513.1 | 3 | 0.01 | 7 |
| Applicant income | 401.5 | 5 | 0.068 | 4 |
| Loan amount | 2141.8 | 1 | 0.016 | 6 |
| Loan purpose | 480.6 | 4 | 0.22 | 1 |
| Lien status | 149.2 | 6 | 0.093 | 2 |
| Loan type | 20.9 | 7 | 0.017 | 5 |

**Panel B: Marginal contribution of each explanatory variable in the SVM model on under-sampled dataset**

Notes. The Table presents the marginal contribution of each explanatory variable in the RF and SVM models in terms of the standard Shapley approach, $\phi(\hat{f}(X_i))$, and linear Shapley-Lorenz approach, $LZ_{d=1}^{X_k}(\hat{y})$, respectively.

## 4.8 Internet Appendix

### 4.8.1 Internet Appendix A

**Table A1: Marginal effects of coefficient estimates of the Logistic Regression model**

| Variable | Marginal effect |
|---|---|
| Applicant race | 0.1084 |
| Applicant gender | 0.0159 |
| Applicant income | 0.0213 |
| Loan amount | -0.0236 |
| Loan purpose | 0.2338 |
| Lien status | -0.1366 |
| Loan type | 0.0725 |

**Panel A: Marginal effects of coefficient estimates on over-sampled dataset**

| Variable | Marginal effect |
|---|---|
| Applicant race | 0.0841 |
| Applicant gender | 0.0218 |
| Applicant income | 0.0255 |
| Loan amount | -0.0406 |
| Loan purpose | 0.2372 |
| Lien status | -0.1320 |
| Loan type | 0.0647 |

**Panel B: Marginal effects of coefficient estimates on under-sampled dataset**

| Variable | Marginal effect |
|---|---|
| Applicant race | 0.1069 |
| Applicant gender | 0.0171 |
| Applicant income | 0.0229 |
| Loan amount | -0.0256 |
| Loan purpose | 0.2326 |
| Lien status | -0.1301 |
| Loan type | 0.0715 |

**Panel C: Marginal effects of coefficient estimates on hybrid-sampled dataset**

Notes. The Table presents the marginal effects of Logistic Regression Models' coefficients. Our dependent variable is *Declined Loan*. All the variables are defined in Table 1.

# Conclusions, Limitations, and Future Work

## 5.1 Introduction

Policymakers, governments, and corporations are increasingly becoming aware of potential environmental risks to business and society. Multiple reports of Intergovernmental Panel on Climate Change emphasize that technological mix in corporations is woefully inadequate for curtailing Global warming. This makes it absolutely imperative to implement a radical change in the mix of technologies used for producing and consuming energy. The World Economic Forum (2019) also identified "failure of climate change mitigation and adaption" as one of the top three risks. Further, BlackRock's Chairman and CEO, Larry Fink, in his letter to the CEOs of the largest companies globally has warned that "climate change is different. Even if only a fraction of the projected impacts is realized, this is a much more structural, long-term crisis. Companies, investors, and governments must prepare for a significant reallocation of capital."[79] Hence, my dissertation examines whether a capital market incentive exists to decarbonise the international economy through a radical change in the mix of technologies that help produce and consume energy, rather than through energy-efficiency improvements of existing carbon-based technologies.

In this dissertation, I also examine whether national culture traits profiling can usefully inform a machine learning alert model to detect money laundering at a globally prominent financial institution. In light of recent literature on the role of culture in corporate misconduct and bank failure (Berger et al., 2021; Liu, 2016; DeBacker et al., 2015; Bame-Aldred et al., 2013), I explore the relevance of several country-specific cultural and institution quality indices vis-à-vis modelling incidence of suspicious money movement within a financial institution.

Finally, I examine whether the feature importance in logistic regression predictive models as indicated by Global Shapley Value and Shapley-Lorenz model-agnostic explainable AI methods align with evidence of feature importance in the underlying models, in the context of real-world financial services bank lending data. My study delivers practitioner-oriented tests and demonstrations on the usefulness of Shapley measures that render opaque but accurate machine learning models useful, in line with the spirit of regulatory supervision governing algorithmic bias and model accountability. The question I address is significant because organizations are increasingly employing machine learning techniques to make decisions that are crucial to our wellbeing, in the hope that such algorithms might counteract human prejudices and inconsistencies. Although such algorithms may yield impressive predictive performances, their obfuscating internal logic may inadvertently perpetuate biases leading to prevention of detection and mitigation of discrimination. Thus, my study provides real-world evidence on the usefulness or otherwise of the explainable AI techniques in uncovering a machine learning model's inter-

[79] Larry Fink, "A Fundamental Reshaping of Finance," BlackRock letter to CEOs (14 January 2020).

nal logic. The explainable AI techniques in revealing whether the algorithms are fair can help organizations mitigate discrimination.

In the next Section, I provide a summary of the thesis, highlight the main findings and literature contributions. In Section 5.3, I discuss the thesis's limitations and present avenues for future research. Finally, in Section 5.4, I provide a summary of the thesis's main findings and conclusions.

## 5.2 Main Findings and Literature Contributions

Innovation productivity is immensely important for firm- and national-level competitiveness in international markets (Porter, 1992). Innovation productivity in enabling environmentally friendly technologies to curtail and reverse environmental degradation can help establish a sustainable market economy around the world (Allen and Yago, 2011; IPCC, 2014). A sustainable market economy besides mitigating market failures can also serve to protect air, water, fisheries, wildlife, and biodiversity. In chapter 2, I raise the question of whether an economic incentive exists for firms to pursue strategies of clean environmentally supportive innovation, as opposed to carbon-emitting dirty innovation activities. In other words, I examine whether firms conducting clean innovation trade at a premium or a discount relative to those which conduct dirty innovation. To answer this question, I avail of capital market price signals to assess the presence and magnitude of economic incentives for clean innovation relative to dirty innovation.

Recent emphasis on clean technologies from fossil fuel-based innovations to curb carbon and other greenhouse gas emissions has inspired both theoretical (Acemoglu et al., 2012)and empirical (Aghion et al., 2016) research in this area. Prior studies foreground evidence that firms may redirect innovation away from fossil fuel towards low carbon technologies, when faced with change in policies and energy prices (Calel and Dechezlepretre, 2016; Popp, 2002; Newell et al., 1999). However, a limitation of existing studies of directed technological change concerns that a multitude of drivers determine companies' decisions to conduct R&D activity in clean or dirty technologies. A complex medley of factors including the relative prices of production factors (Hicks, 1932a; Popp, 2002; Acemoglu et al., 2012), the quality of environmental policy instruments (Johnstone et al., 2010), the extent of market demand, and a path-dependency in knowledge creation (Acemoglu et al., 2012; Aghion et al., 2016) can influence the prospective economic returns of clean and dirty innovation. Most important, many coexisting policies in a given jurisdiction - for example, carbon markets, fuel taxes, energy efficiency standards and renewable energy mandates - make it difficult to measure the overall stringency of environmental regulations affecting the companies. An additional complexity consists in the expected realization of these policies and drivers which determine innovation decisions, rather than current observed realizations. However, these expectations which are

inevitably not directly observed may vary markedly across firms. A major advantage of the approach adopted in chapter 2 is that the stock market evaluation of patented innovation in clean and dirty technologies can reveal the market expectations with respect to the prospective economic performance of the investments which incorporate all their determinants, in particular from policies.

My analysis in chapter 2 mines a global firm-level patent data set, covering 15,217 firms across 12 countries. To construct the clean and dirty innovation measures, I draw patent data from the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office (EPO).[80] Primarily focusing on the patents and citations published by the United States Patent and Trademark Office (USPTO), in line with existing studies, I also analyse the patents and citations published by the EPO to account for robustness. The baseline empirical analysis focuses on a sample of USPTO published patents and citations filed by 15,217 firms belonging to the top 12 leader countries in clean innovation during 1995-2012.[81] Further, I employ firm-level data from Worldscope and Datastream databases. To match the firm-level data with patent data, I employ Bureau van Dijk's matching algorithm provided under the "IP" bundle of the Orbis database.

After matching the firm-level data with patent-level data, I create the innovation variables. While these variables are inspired by prior literature (Deng et al., 1999; Chan et al., 2001; Hall et al., 2005; Hirshleifer et al., 2013), the chief novelty of my study consists in disaggregating these into 'clean,' 'dirty,' and 'other' components. I associate 'clean' innovation with renewable energy generation, electric vehicles, and energy efficiency technologies in the buildings sector; 'dirty' innovation with fossil-based energy generation and ground transportation. I rely on prior literature to disaggregate the innovation measures into clean and dirty innovation categories. To identify clean innovation, I rely on the previous work of the OECD Environment Directorate that lists the patent classification codes for clean technologies.[82] I follow Noailly and Smeets (2015) and Aghion et al. (2016) to identify dirty technologies.[83]

In chapter 2, I adapt a firm-level market-value function (Griliches, 1981; Hall et al., 2005; Hall and Oriani, 2006) and Fama-MacBeth regressions (Fama and MacBeth, 1973). In the light of these models, I draw inferences on how the innovation variables influence the firm's Tobin's Q.

[80] In reporting the name of applicants, the patent database allows me to match clean and dirty patents with distinct patent holders. Further, the database in including information on patent citations, allows me to address the well-known issue of heterogeneity in patent value.

[81] The top 12 clean innovation producing countries in descending order are: Japan, USA, Korea, Germany, Taiwan, France, Denmark, Netherlands, Canada, Sweden, Finland, and Great Britain. Patenting at the USPTO in clean and dirty technologies becomes miniscule beyond these top 12 countries.

[82] I examine areas of clean patenting activity related to energy generation from renewable and non-fossil sources (wind, solar, hydro, marine, biomass, geothermal and energy from waste), combustion technologies with mitigation potential (for example, combined heat and power), other technologies with potential contribution to emissions mitigation (in particular energy storage), electric and hybrid vehicles and energy conservation in buildings.

[83] I employ patent classification codes for dirty technologies in electricity generation industry from Noailly and Smeets (2015) and in automobile industry from Aghion et al. (2016), respectively.

First, I verify the capital market value accorded to generic innovation productivity (Deng et al., 1999; Chan et al., 2001) and innovation efficiency (Hirshleifer et al., 2013). This work serves to extend, in the international arena, the non-linear least squares regression model findings in Hall et al. (2005).[84] Then, to determine the expected economic performance of 'clean' and 'dirty' investment, I disaggregate the innovation variables into clean, dirty, and other categories. I find that an additional clean patent, per million dollars of book value, is associated with an increment of 3.77% in Tobin's Q. I also find that generating a citation on a clean patent, per million dollars of book value, is associated with an increment of 1.27% in Tobin's Q. I also note that the comparable efficiency of R&D investments, in generating dirty patents, reduces the market value of the firm to the tune of 0.97% of its economic value. My main finding is, thus, that 'clean' innovation is associated with an economically important and positive Tobin's Q relation, especially relative to the inferred association with dirty innovation. Further, I adopt a wide variety of complementary and state-of-the-art testing procedures to investigate whether clean innovation is associated with firm value. These tests are based on a variety of dimensions:

1. Following Hirshleifer et al. (2013), I test if the findings are invariant to the Fama-Macbeth two-step regression estimator (Fama and MacBeth, 1973).

2. I test if the results are robust when I restrict the sample to include only those firms that conduct both clean and dirty innovation.

3. I test if the results can be accounted for by including emerging technology innovation in the baseline regressions.

4. I check the sensitivity of the results to include a range of firm traits from the accounting-based asset pricing literature (Ohlson, 1989, 1995; Hirshleifer et al., 2013).

5. I conduct a Heckman two-stage analysis (Heckman, 1979) to account for sample selection concerns.

6. Finally, I test if the main findings hold when I examine European patents, as opposed to United States patents.

Across the conducted tests, I show evidence of the importance of clean innovation (but not dirty innovation) for the equity market's indication of firm value.

Prior literature foreground evidence that capital markets can create financial and reputational incentives for pollution control in both developed and emerging market economies (Gupta and Goldar, 2005; Hamilton, 1995; Dasgupta et al., 2001). Further, prior literature finds a positive association between firm-level eco-efficiency with operating performance and market value

[84]The initial findings corroborate a large body of research that provide compelling evidence of the patent productivity of R&D and the citations received by these patents having a statistically and economically significant positive impact on firms' market value (Griliches, 1981; Chan et al., 2001; Eberhart et al., 2004).

(Guenster et al., 2011; Ziegler et al., 2007; Von Arx and Ziegler, 2014). However, these studies suffer from several limitations including the problematic small samples and the lack of objective environmental performance criteria. In my thesis, I have tried to overcome the limitations of prior studies. For instance, instead of relying on subjective analysis to characterize environmental performance, I study the documented environmental patenting activity and the efficiency of this patenting activity of publicly traded firms around the world. Additionally, unlike prior literature, my study examines the critically important performance criterion of environmentally friendly patented innovation with a view to improving the mix of technologies used to produce and consume energy (IPCC, 2014).

While my study does not aim to establish the underlying mechanism that can account for a clean or dirty innovation premium, it, nonetheless, highlights several competing or complementary explanations that can drive the existence of a clean innovation premium. Therefore, empirically investigating the drivers behind the clean innovation premium uncovered in the chapter provides an important avenue for research.

In chapter 3, I examine the utility and ethics of incorporating national culture profiling in bank-level machine-learning informed alert models relating to financial malfeasance. Specifically, I test to establish the utility of national culture traits informing a machine learning alert model for detecting money laundering at a globally prominent financial institution. To address this question, I employ a major global financial institution's large proprietary dataset containing cross-border wire transactions made during 2009-2018. Those wires that the institution's designated investigative team flagged as 'suspicious activities' can be regarded as precursors to money laundering. I further collate the novel proprietorial customer and account level cross-border wire transfer bank client data with country-specific culture (Hofstede's cultural dimensions) and institution quality indices (Corruption Perception Index; Financial Secrecy Index). Further, the proprietorial dataset provides a clearly labelled response variable (Issue Case). I, therefore, employ supervised learning techniques such as logistic regressions, random forest, gradient boosted machines, and support vector machines to detect money-laundering at the financial institution. Employing the said machine learning techniques together with corrections for data imbalance, the results reflect the strength of national culture dimensions in formulating anti-money laundering (AML) predictions. I find that for both corporate-related and combined alerts, the individuality rating of both the customer's residence country $(IDV_R)$ and country of wire origination/destination $(IDV_W)$ are of paramount importance. This is followed by the corruption perception score of the country of wire origination/destination $(CPI_W)$ and the customer's residence country $(CPI_R)$ for the corporate-related alerts; and $CPI_W$ and the financial secrecy score of the customer's resident country $(FSI_R)$ for the combined alerts. For people-related alerts, the corruption perception score for the country of wire origination/destination $(CPI_W)$ and the financial secrecy score of the resident country $(FSI_R)$ are the two most important features, followed by the $CPI_R$ and $IDV_R$.

Further, the introduction of these variables complements the institution's own account- and transaction-level data, considering that the inclusion of these predictors as an added layer of security enhances the performance of the models. For corporate-related alerts, the county-level features that rank among the top five include the individuality rating of the customer's country of residence ($IDV_R$), individuality rating of the country of wire origination/destination ($IDV_W$), and the uncertainty avoidance cultural trait of the customer's residence country ($UAI_R$). I further find that the power-distance index score of the customer's residence country ($PDI_R$) informs the customer's predilections for committing financial misconduct. For people-related alerts, the individuality score of the customer's residence country ($IDV_R$), corruption perception score of the country of wire origination/destination ($CPI_W$), and financial secrecy score of the customer's country of residence ($FSI_R$) are the most important county-level features that rank among the top ten features. These results provide evidence of the usefulness of culture traits of customers for detecting both corporate and individual malfeasance.

My findings provide practical implications for the financial services sector in terms of AML compliance and prevention strategy. Confirming the conduciveness of machine learning in incorporating national culture, the findings also contribute to the extensive literature that ascribes values to ethicality and discernment constituting distinct national traits.

Finally, chapter 4 examines whether the feature importance in logistic regression predictive models as indicated by Global Shapley Value and Shapley-Lorenz model-agnostic explainable AI methods align with evidence of feature importance in the underlying models, in the context of real-world financial services bank lending data. Scholarship confirms the validity of Shapley value-based model-agnostic explainable AI methods on simulated datasets (Strumbelj and Kononenko, 2010; Strumbelj and Kononenko, 2014; Aas et al., 2021).[85] However, evidence of their usefulness on real-world datasets is scant, particularly in respect to impactful financial decisions. The methodology adopted in chapter 4 involves the estimation of tractable and transparent machine learning model, logistic regression, in mortgage lending data to discern the relative importance of predictive features. It then deploys Global Shapley Value and Shapley-Lorenz explainable AI methods to test if their insight concerning feature importance is in line with that of the logistic regression model. Finally, it examines whether these methods can enable financial institutions to select an opaque creditworthiness assessment model which blends out-of-sample performance with ethical considerations.

Prior literature finds evidence of racial discrimination in both in-person and algorithmic lending in the US (Bartlett et al., 2021; Butler et al., 2020; Blanchflower et al., 2003; Munnell et

[85]For instance, Strumbelj and Kononenko (2010; 2014) use Shapley value-based feature importance measurements, via their approximation method, and show accurate results across various data generating processes. They use various learning algorithms such as decision trees, naïve bayes, support vector machines, multi-layer perceptron artificial neural networks, random forest, logistic regression and ADaBoost to evaluate and validate their approximation method. They further evaluate their method's usefulness on a real-world oncology dataset.

al., 1996; Black et al., 1978). Thus, I test if Global Shapley Value and Shapley-Lorenz XAI methods can uncover algorithmic injustice in the bank lending space. Further, 157,269 loan applications from Home Mortgage Disclosure Act's (HMDA) website made in New York during 2017 is examined. I first deploy a logistic regression model and show evidence consistent with racial discrimination. I then test if the said XAI methods give insight consistent with the logistic regression model. Accordingly, I find that these XAI methods establish the prevalence of racial discrimination as a paramount factor. In revealing that the XAI methods uncover racial discrimination, the analysis confirms their validity in respect to the logistic regression model, and in real-world datasets. This chapter also shows how financial institutions can derive accurate and accountable decisions, in the context of racial discrimination and opaque creditworthiness models.

## 5.3 Limitations and Future Work

Chapter 2 investigates whether a clean innovation premium, consistent with the objective for a long-term decarbonization of the international economy is feasible. My purpose in this chapter is not to establish the underlying mechanism that can account for a clean or a dirty innovation premium. However, I seek to investigate whether a clean or dirty innovation premium informs the data. Life-cycle argument is one possible mechanism that could drive clean innovation premium. Early-stage life-cycle technology can be associated with potential for high growth albeit with high risk. If initially assets are valued higher than their replacement cost, competition in the marketplace will erode this markup over time (Tobin, 1969). This life-cycle argument leading to smaller effects of incremental patenting on Tobin's q for a given technology over time can potentially account for my main finding on clean innovation premium. Thus, I undertake exploratory and tentative work to investigate the importance of the life-cycle phase of the clean and dirty patented technology. Specifically, I conduct clean innovation premium tests (1) in relation to an industry sector, Drugs, where growth rates over time in clean and dirty technology patents are comparable (I tentatively assume that this indicates that clean and dirty technologies are comparably mature) and (2) related to new emergent technologies (which are presumably at the early stage of their technology life cycle). I note that these findings provide some initial (albeit mixed) evidence of the importance of life-cycle argument in accounting for a clean innovation premium. My study, however, highlights several competing or complementary explanations that can drive the existence of a clean innovation premium. Therefore, empirically investigating the drivers behind the clean innovation premium uncovered in the chapter provides an important avenue for research.

My models in investigating only simplified 'hedonic' market value equations (using non-linear least squares and Fama-Macbeth estimators to model firm value on sets of covariates) do not address the deeper dynamic forces that impact the correlation between successful patent evaluation and a corresponding stock market evaluation (e.g., see Pakes, 1985). Thus, one promising

avenue for future research consists in investigating the deeper dynamic forces at work (Pakes, 1985).

Further, certain asymmetry underlies the way clean and dirty innovations are defined. Clean technologies encompass a larger set of sectors than dirty technologies. The main challenge I encountered stemmed from the difficulty associated with systematically identifying dirty technologies corresponding to all clean technologies in the database. While it is easy to identify combustion engines and coal-fired electricity production technologies, it is rather difficult to identify building materials that are either less or not at all energy-efficient compared to the technologies associated with energy conservation in buildings defined as clean. To account for this asymmetry, I have incorporated the following:

1. I include sector fixed effects in all the models to avoid the results being driven by industry differences.

2. I test for the presence of clean/dirty innovation premium for a variety of subsamples. I also test those sample of firms conducting both clean and dirty innovation. By testing on this particular sample, I neatly eliminate firms that potentially innovate only in clean technologies, and thereby mitigate overestimation of the clean innovation premium.

3. Firms producing clean innovation whose dirty equivalents I have not observed may also, potentially, be innovating in dirty technologies. These dirty innovation patents are, however, included in the 'other' patent category, which systematically informs the models. I find that clean innovation is associated with a larger Tobin's q compared with both dirty and 'other' technologies.

To compellingly address the issue of asymmetry in how clean and dirty innovation are defined, I could, however, primarily delimit my analysis to only the energy and automobile sector. While this is feasible, it would also render the study much less general, considering many clean innovations necessary to decarbonize the economy will have to be sourced from other sectors, in particular the buildings and the manufacturing sectors. I leave this avenue of research for future work. Secondarily, I could identify dirty innovation for those sectors where I observed clean technologies. This too is left for future research.

A further limitation in Chapter 2 arises from employing patents as a proxy for firms' innovativeness. The information content, i.e., the evaluation of patents in regard to measuring firm-level innovation, can markedly vary across industry sectors, jurisdictions and, indeed, time. Patent measures, even across individual firms, can arguably be viewed as a noisy proxy for the development of a firm's technology.

That said, a focus on inventions belonging to a particular technological class for a single industry, for instance, would forgo an analysis of the economically important question that this

study seeks to address: whether a capital market incentive exists to decarbonize the international economy. Predictably previous empirical studies have largely used patent counts as an indicator of innovation often neglecting to account for the heterogeneity across settings and industries.

Griliches (1998) acknowledges this cadre of difficulty: 'not all inventions are patentable, not all inventions are patented, and the inventions that are patented differ greatly in quality.' He recommends that the first two problems can be tackled by industry dummy variables (or by limiting the analysis to a particular industry sector); the third problem can be addressed, despite heterogeneity in the data, by invoking the 'law of large numbers' (the economic significance of a patent can be usefully thought of as a random number variable with a probability distribution).

Keeping the elaborations of Griliches (1998) in mind, I attempt to address the issues raised by

1. Specifying industry and time dummies, and also country dummies (in Internet Appendix H) in the regression specifications. As Griliches (1998) indicates this can serve to alleviate the concern of differing effects across industry sectors and over time.

2. Invoking the law of large numbers. I examine a global patent data set across 15,000 firms in 12 countries. It is the largest data set as far as my knowledge goes that examines the economic value of a patent.

I employ adjusted patent citations in my models (Hirshleifer et al., 2013; Pandit et al., 2011; Gu, 2005). My measures of citation productivity and efficiency comprise citation of the year, technological fields, and grant year account for the citation propensity. It is widely accepted that citations of a firm's patents indicate the technological and economic significance of the innovation (Hall et al., 2005; Hirshleifer et al., 2013). Given that even this set of well-informed adjustments cannot fully alleviate the concern, a finding borne out by the referenced literature, my inferences too suffer this bias.

In Chapter 3, using proprietorial wire money transfer data of a globally prominent financial institution, I evaluate national culture profiling in formulating machine learning models to counter money laundering. To discern the importance of national culture in both parsimonious and heavily parameterised machine learning models which comprise a wealth of customer and account level data, I use a weighted average of the Gini index variable importance scores of the random forest and gradient boosted machine models. I acknowledge that the weighted average measure is not an accurate measure to rank the variables, since the variable importance methods incorporated in my study are model-specific. Hence to alleviate this shortcoming, I propose to employ model-agnostic explainable AI methods in my future work. These methods will enable me to compare the results of various models, even as they inform the interpretability of models for which model-specific interpretation methods do not exist. I leave it to future work to employ explainable model-agnostic AI methods to discern the importance of national culture in

machine learning models. This will help gain insights into the precise importance of national culture traits, relative to customers' account and transaction traits, in the performance of these models.

Chapter 4 provides real-world evidence on the usefulness or otherwise of the Global Shapley Value and Shapley-Lorenz explainable AI methods in uncovering racial discrimination in the bank lending space. I employ only the loan applications data provided by HMDA without augmenting it with relevant borrower characteristics. I acknowledge that the logistic regression models' results may be relaying the effect of unobserved financial traits of the applicants. However, in my study I raise the question whether XAI methods could give results consistent with the logistic regression model. That is, my objective is in testing whether the XAI methods could uncover racial discrimination rather than its incidence in New York lending decisions. Hence, the purpose of my study consists in demonstrating the usefulness of explainable model-agnostic methods in algorithmic lending. In my future research, I plan to extend my dataset to include the financial traits of the applicants to study the usefulness or otherwise of the said XAI methods in uncovering racial discrimination in the New York bank lending space. With the enlarged dataset, I also plan to test the usefulness of the XAI in the US bank lending space, as a whole.

## 5.4 Summary and Conclusions

I begin this chapter by foregrounding the thesis's major findings and contributions. Further, I discuss the thesis's limitations and present a few avenues for future research.

This dissertation through rigorous state-of-the-art statistical techniques addresses pertinent and timely research questions that carry enormous social impact. The thesis consists of three essays that address important social questions on climate change and ethical AI in the financial economics space.

Chapter 2 that deals with environmental finance address whether an economic incentive obtains for firms to pursue strategies of clean environmentally supportive innovation, as opposed to carbon-emitting dirty innovation. This question is informed by several reports of Intergovernmental Panel on Climate Change (IPCC) which indicate that stabilizing global carbon emissions by 2050 will require a 60% reduction in the carbon intensity of global GDP, compared with a business-as-usual scenario. Hence, the chapter is motivated by whether a capital market incentive exists to decarbonise the international economy through a radical change in the mix of technologies that help produce and consume energy, rather than through energy-efficiency improvements of extant carbon-based technologies. Using a global patent dataset covering over 15,000 firms across 12 countries, chapter 2 uncovers strong and robust evidence that the stock market recognizes the value of clean innovation and innovation efficiency and accords higher valuations to those firms that engage in successful clean research and development activities.

The results are substantively invariant across innovation measurement, model specifications, estimators adopted, select sub-samples of firms, and the United States and European patent offices.

Chapters 3 and 4 that discuss financial data science assess the utility and ethics of incorporating national culture profiling in bank-level machine-learning informed alert models relating to financial malfeasance and tests state-of-the-art explainable AI techniques to uncover algorithmic injustice in the bank lending space, respectively.

In chapter 3, I test to establish the utility of national culture traits informing a machine learning alert model for detecting money laundering at a globally prominent financial institution. My findings reflect the strength of national culture dimensions in formulating anti-money laundering predictions. Further, the national culture variables complement the institution's own account- and transaction-level data, considering that the inclusion of these predictors as an additional layer of security enhances the performance of the models. These findings provide practical implications for the financial services sector in terms of AML compliance and prevention strategy. Confirming the conduciveness of machine learning in incorporating national culture, the findings also contribute to the extensive literature that ascribes values to ethicality and discernment constituting distinct national traits. This chapter further provides the first description of the ethics associated with employing national culture profiles in machine-learning to counter money laundering.

Finally, in chapter 4, I test the validity of Global Shapley Value and Shapley-Lorenz model-agnostic explainable AI methods on a real-world finance dataset. Regulators and various national agencies in the US (USACM, 2017; OSTP Report, 2016) and Europe (European Commission, 2019; France, 2018a and 2018b) are increasingly recognising the importance of algorithmic transparency and accountability. They encourage the use of Machine Learning models that ensure high predictive performance as well as interpretability. For instance, the European Commission emphasises the importance of research in explainable AI systems to render transparent and accountable high performance machine learning models with a view to ensuring the protection of customer rights. Although black-box models may yield impressive predictive performance, their obfuscating internal logic may inadvertently perpetuate biases leading to prevention of detection and mitigation of discrimination. In revealing the importance of features that determine the machine learning models' decisions, the state-of-the-art explainable model-agnostic methods can uncover algorithmic biases and, thereby allow institutions to employ "fairness" techniques for rectifying the error. Prior literature finds evidence of racial discrimination in both in-person and algorithmic lending in the US (Bartlett et al., 2021; Butler et al., 2020; Blanchflower et al., 2003; Munnell et al., 1996; Black et al., 1978). Thus, I test if Global Shapley Value and Shapley-Lorenz XAI methods can uncover algorithmic injustice in the bank lending space. I first deploy a logistic regression model and show evidence consistent

with racial discrimination. I then test if the said XAI methods give insight consistent with the logistic regression model. Accordingly, I find that these XAI methods establish the prevalence of racial discrimination as a paramount factor. In revealing that the XAI methods uncover racial discrimination, the analysis confirms their validity in respect to the logistic regression model, and in real-world datasets. The chapter also shows how financial institutions can derive accurate and accountable decisions, in the context of racial discrimination and opaque creditworthiness models.

# References


Aas, K., M. Jullum, and A. Løland (2021). Explaining individual predictions when features are dependent: More accurate approximations to shapley values. *Artificial Intelligence 298*, 103502.

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012). The environment and directed technical change. *American economic review 102*(1), 131–66.

Adomavicius, G. and A. Tuzhilin (2001). Using data mining methods to build customer profiles. *Computer 34*(2), 74–82.

Aggarwal, R., J. E. Goodell, and J. W. Goodell (2014). Culture, gender, and gmat scores: Implications for corporate ethics. *Journal of Business Ethics 123*(1), 125–143.

Aggarwal, R. and J. W. Goodell (2009). Markets and institutions in financial intermediation: National characteristics as determinants. *Journal of Banking & Finance 33*(10), 1770–1780.

Aggarwal, R. and J. W. Goodell (2013). Political-economy of pension plans: Impact of institutions, gender, and culture. *Journal of Banking & Finance 37*(6), 1860–1879.

Aggarwal, R., J. W. Goodell, and L. J. Selleck (2015). Lending to women in microfinance: Role of social trust. *International Business Review 24*(1), 55–65.

Aghion, P., A. Dechezleprêtre, D. Hemous, R. Martin, and J. Van Reenen (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy 124*(1), 1–51.

Allen, F. and G. Yago (2011). Environmental finance: Innovating to save the planet. *Journal of Applied Corporate Finance 23*(3), 99–111.

Alter, A. L. and J. M. Darley (2009). When the association between appearance and outcome contaminates social judgment: A bidirectional model linking group homogeneity and collective treatment. *Journal of Personality and Social Psychology 97*(5), 776.

Angwin, J., J. Larson, S. Mattu, and L. Kirchner (2016). Machine bias: There's software used across the country to predict future criminals. and it's biased against blacks. propublica, may 23.

Angwin, J., T. Parris Jr, and S. Mattu (2016). Breaking the black box: What facebook knows about you. *ProPublica*.

Arce, D. G. and M. C. Gentile (2015). Giving voice to values as a leverage point in business ethics education. *Journal of Business Ethics 131*(3), 535–542.

Armstrong, R. W. (1996). The relationship between culture and perception of ethical problems in international marketing. *Journal of Business Ethics 15*(11), 1199–1208.

Ballardini, F., A. Malipiero, R. Oriani, M. Sobrero, and A. Zammit (2005). Do stock markets value innovation? A meta-analysis. *A Meta-Analysis (January 2005)*.

Bame-Aldred, C. W., J. B. Cullen, K. D. Martin, and K. P. Parboteeah (2013). National culture and firm-level tax evasion. *Journal of Business Research 66*(3), 390–396.

Barth, M. E., W. H. Beaver, and W. R. Landsman (1998). Relative valuation roles of equity book value and net income as a function of financial health. *Journal of Accounting and Economics 25*(1), 1–34.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2021). Consumer-lending discrimination in the fintech era. *Journal of Financial Economics*.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2022). Consumer-lending discrimination in the fintech era. *Journal of Financial Economics 143*(1), 30–56.

Batista, G. E., R. C. Prati, and M. C. Monard (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD explorations newsletter 6*(1), 20–29.

Baxamusa, M. and A. Jalal (2014). Does religion affect capital structure? *Research in International Business and Finance 31*, 112–131.

Berger, A. N., X. Li, C. Morris, and R. A. Roman (2019). The effects of cultural values on bank failures around the world. *Journal of Financial and Quantitative Analysis (JFQA), Forthcoming*.

Berger, A. N., X. Li, C. S. Morris, and R. A. Roman (2021). The effects of cultural values on bank failures around the world. *Journal of Financial and Quantitative Analysis 56*(3), 945–993.

Berk, R. (2017). An impact assessment of machine learning risk forecasts on parole board decisions and recidivism. *Journal of Experimental Criminology 13*(2), 193–216.

Biais, B., D. Hilton, K. Mazurier, and S. Pouget (2005). Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *The Review of economic studies 72*(2), 287–312.

Black, H., R. L. Schweitzer, and L. Mandell (1978). Discrimination in mortgage lending. *The American Economic Review 68*(2), 186–191.

Black, H. A., T. P. Boehm, and R. P. DeGennaro (2003). Is there discrimination in mortgage pricing? the case of overages. *Journal of Banking & Finance 27*(6), 1139–1165.

Blanchflower, D. G., P. B. Levine, and D. J. Zimmerman (2003). Discrimination in the small-business credit market. *Review of Economics and Statistics 85*(4), 930–943.

Blundell, R., R. Griffith, and J. Van Reenen (1999). Market share, market value and innovation in a panel of British manufacturing firms. *The Review of Economic Studies 66*(3), 529–554.

Botta, E. and T. Koźluk (2014). Measuring environmental policy stringency in OECD countries.

Bowman, E. H. and M. Haire (1975). A strategic posture toward corporate social responsibility. *California management review 18*(2), 49–58.

Brekke, K. A. and K. Nyborg (2008). Attracting responsible employees: Green production as labor market screening. *Resource and Energy Economics 30*(4), 509–526.

Brennan, M. J. (1991). A perspective on accounting and stock prices. *The Accounting Review 66*(1), 67–79.

Brewer, M. B. and A. S. Harasty (1996). Seeing groups as entities: The role of perceiver motivation.

Buhmann, A., J. Paßmann, and C. Fieseler (2019). Managing algorithmic accountability: Balancing reputational concerns, engagement strategies, and the potential of rational discourse. *Journal of Business Ethics*, 1–16.

Butler, A. W., E. J. Mayer, and J. Weston (2020). Racial discrimination in the auto loan market. *Available at SSRN 3301009*.

Calel, R. and A. Dechezlepretre (2016). Environmental policy and directed technological change: Evidence from the european carbon market. *Review of economics and statistics 98*(1), 173–191.

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science 3*(1), 14.

Castro, J., D. Gómez, and J. Tejada (2009). Polynomial calculation of the shapley value based on sampling. *Computers & Operations Research 36*(5), 1726–1730.

Chan, L. K., J. Lakonishok, and T. Sougiannis (2001). The stock market valuation of research and development expenditures. *The Journal of Finance 56*(6), 2431–2456.

Chen, C. C., M. W. Peng, and P. A. Saparito (2002). Individualism, collectivism, and opportunism: A cultural perspective on transaction cost economics. *Journal of Management 28*(4), 567–583.

Cheng, P., Z. Lin, and Y. Liu (2015). Racial discrepancy in mortgage interest rates. *The Journal of Real Estate Finance and Economics 51*(1), 101–120.

Chui, A. C., A. E. Lloyd, and C. C. Kwok (2002). The determination of capital structure: is national culture a missing piece to the puzzle? *Journal of international business studies 33*(1), 99–127.

Chui, A. C., S. Titman, and K. J. Wei (2010). Individualism and momentum around the world. *The Journal of Finance 65*(1), 361–392.

Cobham, A., P. Janskỳ, and M. Meinzer (2015). The financial secrecy index: Shedding new light on the geography of secrecy. *Economic Geography 91*(3), 281–303.

Colombo, E. and M. Pelagatti (2020). Statistical learning and exchange rate forecasting. *International Journal of Forecasting 36*(4), 1260–1289.

Cook, J. (2008). Ethics of data mining. In *Information Security and Ethics: Concepts, Methodologies, Tools, and Applications*, pp. 211–217. IGI Global.

Corbett, C. and S. Muthulingam (2008). An empirical investigation of the depth of adoption of the LEED green building standards. Technical report, Working Paper, UCLA.

Cornett, M. M., J. J. McNutt, P. E. Strahan, and H. Tehranian (2011). Liquidity risk management and credit supply in the financial crisis. *Journal of financial economics 101*(2), 297–312.

Courchane, M. and D. Nickerson (1997). Discrimination resulting from overage practices. In *Discrimination in Financial Services*, pp. 133–151. Springer.

Cullen, J. B., K. P. Parboteeah, and M. Hoegl (2004). Cross-national differences in managers' willingness to justify ethically suspect behaviors: A test of institutional anomie theory. *Academy of Management Journal 47*(3), 411–421.

Czarnitzki, D., B. H. Hall, and R. Oriani (2006). The market valuation of knowledge assets in US and European firms. *The Management of Intellectual Property, Cheltenham Glos*, 111–131.

Dasgupta, N., M. R. Banaji, and R. P. Abelson (1999). Group entitativity and group perception: Associations between physical features and psychological judgment. *Journal of Personality and Social Psychology 77*(5), 991.

Dasgupta, S., B. Laplante, and N. Mamingi (2001). Pollution and capital markets in developing countries. *Journal of Environmental Economics and management 42*(3), 310–335.

Davidson, R., A. Dey, and A. Smith (2015). Executives'"off-the-job" behavior, corporate culture, and financial reporting risk. *Journal of Financial Economics 117*(1), 5–28.

Davis, J. H. and J. A. Ruhe (2003). Perceptions of country corruption: Antecedents and outcomes. *Journal of Business Ethics 43*(4), 275–288.

DeBacker, J., B. T. Heim, and A. Tran (2015). Importing corruption culture from overseas: Evidence from corporate tax evasion in the united states. *Journal of Financial Economics 117*(1), 122–138.

Deng, Z., B. Lev, and F. Narin (1999). Science and technology as predictors of stock performance. *Financial Analysts Journal 55*(3), 20–32.

Dimmock, S. G. and W. C. Gerken (2012). Predicting fraud by investment managers. *Journal of Financial Economics 105*(1), 153–173.

Donaldson, T. and T. W. Dunfee (1994). Toward a unified conception of business ethics: Integrative social contracts theory. *Academy of management review 19*(2), 252–284.

Dowell, G., S. Hart, and B. Yeung (2000). Do corporate global environmental standards create or destroy market value? *Management science 46*(8), 1059–1074.

Drehmann, M. and N. Tarashev (2013). Measuring the systemic importance of interconnected banks. *Journal of Financial Intermediation 22*(4), 586–607.

Eberhart, A. C., W. F. Maxwell, and A. R. Siddique (2004). An examination of long-term abnormal stock returns and operating performance following R&D increases. *The Journal of Finance 59*(2), 623–650.

Efendi, J., A. Srivastava, and E. P. Swanson (2007). Why do corporate managers misstate financial statements? the role of option compensation and other factors. *Journal of financial economics 85*(3), 667–708.

Ellerman, A. D., C. Marcantonini, and A. Zaklan (2014). The eu ets: Eight years and counting. *Robert Schuman Centre for Advanced Studies Research Paper* (2014/04).

European Commission, . (2019). Ethics guidelines for trustworthy ai. european commission, https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

Fama, E. F. and J. D. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy 81*(3), 607–636.

FATF (2021). Opportunities and challenges of new technologies for aml/cft. *FATF, Paris, France*.

FATF and Egmont Group (2020). Trade-based money laundering: Risk indicators. *FATF, Paris, France*.

FATF, F. A. T. F. (1999). Money laundering. *Policy Brief July 1999*.

Financial Stability Board (2018). Strengthening governance frameworks to mitigate misconduct risk: A toolkit for firms and supervisors.

Fischer, R. and A. Chalmers (2008). Is optimism universal? a meta-analytical investigation of optimism levels across 22 nations. *Personality and Individual Differences 45*(5), 378–382.

Fischer, R., C.-M. Vauclair, J. R. Fontaine, and S. H. Schwartz (2010). Are individual-level and country-level value structures different? testing hofstede's legacy with the schwartz value survey. *Journal of cross-cultural psychology 41*(2), 135–151.

Fisher-Vanden, K. and K. S. Thorburn (2011). Voluntary corporate environmental initiatives and shareholder wealth. *Journal of Environmental Economics and management 62*(3), 430–445.

Fisman, R. and E. Miguel (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political economy 115*(6), 1020–1048.

France (2018). AI for humanity: French strategy for artificial intelligence. president of the french republic.

Freund, Y., R. E. Schapire, et al. (1996). Experiments with a new boosting algorithm. In *icml*, Volume 96, pp. 148–156. Citeseer.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther (2018). Predictably unequal? the effects of machine learning on credit markets. *The Effects of Machine Learning on Credit Markets (November 6, 2018)*.

Gaganis, C., I. Hasan, P. Papadimitri, and M. Tasiou (2019). National culture and risk-taking: Evidence from the insurance industry. *Journal of Business Research 97*, 104–116.

Getz, K. A. and R. J. Volkema (2001). Culture, perceived corruption, and economics: A model of predictors and outcomes. *Business & society 40*(1), 7–30.

Ghent, A. C., R. Hernandez-Murillo, and M. T. Owyang (2014). Differences in subprime loan pricing across races and neighborhoods. *Regional Science and Urban Economics 48*, 199–215.

Giudici, P. and E. Raffinetti (2020). Lorenz model selection. *Journal of Classification 37*(3), 754–768.

Giudici, P. and E. Raffinetti (2021). Shapley-lorenz explainable artificial intelligence. *Expert Systems with Applications 167*, 114104.

Goodell, J. W. (2019). Comparing normative institutionalism with intended rationality in cultural-finance research. *International Review of Financial Analysis 62*, 124–134.

Grandi, A., B. H. Hall, and R. Oriani (2009). R&D and financial investors. *Evaluation and Performance Measurement of Research and Development, Cheltenham, UK: Edward Elgar*, 143–165.

Griffin, J. M., S. Kruger, and G. Maturana (2017). Do personal ethics influence corporate ethics. *Available at SSRN: https://ssrn.com/abstract=2745062*.

Griliches, Z. (1981). Market value, R&D, and patents. *Economics letters 7*(2), 183–187.

Griliches, Z. (1998). Introduction to" r&d and productivity: The econometric evidence". In *R&D and productivity: The econometric evidence*, pp. 1–14. University of Chicago Press.

Gu, F. (2005). Innovation, future earnings, and market efficiency. *Journal of Accounting, Auditing & Finance 20*(4), 385–418.

Guenster, N., R. Bauer, J. Derwall, and K. Koedijk (2011). The economic value of corporate eco-efficiency. *European Financial Management 17*(4), 679–704.

Gupta, S. and B. Goldar (2005). Do stock markets penalize environment-unfriendly behaviour? Evidence from India. *Ecological economics 52*(1), 81–95.

Hall, B. (2000). Innovation and market value,[w:] Productivity, innovation and economic performance, edsd. R. Barro, G. Mason, M. O'Mahoney.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). Market value and patent citations. *RAND Journal of economics*, 16–38.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001). The NBER patent citation data file: Lessons, insights and methodological tools. Technical report, National Bureau of Economic Research.

Hall, B. H. and R. Oriani (2006). Does the market value R&D investment by European firms? Evidence from a panel of manufacturing firms in France, Germany, and Italy. *International Journal of Industrial Organization 24*(5), 971–993.

Hamilton, J. T. (1995). Pollution as news: Media and stock market reactions to the toxics release inventory data. *Journal of environmental economics and management 28*(1), 98–113.

Harhoff, D., F. M. Scherer, and K. Vopel (2003). Citations, family size, opposition and the value of patent rights. *Research policy 32*(8), 1343–1363.

Harris, D. A. (1996). Driving while black and all other traffic offenses: The supreme court and pretextual traffic stops. *J. Crim. L. & Criminology 87*, 544.

Hartney, C. (2009). *Created equal: Racial and ethnic disparities in the US criminal justice system*. National Council on Crime and Delinquency.

Hassan, O. A. and G. Giorgioni (2015). Analyst coverage, corruption and financial secrecy: a multi-country study. *Corruption and Financial Secrecy: A Multi-Country Study (February 18, 2015)*.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.

Heine, S. J. (2003). An exploration of cultural variation in self-enhancing and self-improving motivations.

Heine, S. J., D. R. Lehman, H. R. Markus, and S. Kitayama (1999). Is there a universal need for positive self-regard? *Psychological review 106*(4), 766.

Heinkel, R., A. Kraus, and J. Zechner (2001). The effect of green investment on corporate behavior. *Journal of financial and quantitative analysis 36*(4), 431–449.

Hicks, J. R. (1932a). Marginal productivity and the principle of variation. *Economica* (35), 79–88.

Hicks, J. R. (1932b). The theory of wages. *London: Macmillan*.

Hirshleifer, D., P.-H. Hsu, and D. Li (2013). Innovative efficiency and stock returns. *Journal of Financial Economics 107*(3), 632–654.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications.

Houqe, N., R. M. Monem, M. Tareq, and T. van Zijl (2015). Secrecy and mandatory ifrs adoption on earnings quality.

Husted, B. W. (2000). The impact of national culture on software piracy. *Journal of Business Ethics 26*(3), 197–211.

IMF, I. M. F. (2021). The imf and the fight against money laundering and the financing of terrorism. *Policy Brief July 2021*.

IPCC (2014). Climate change 2014: Synthesis report. Contribution of Working Groups I. *II and III to the Fifth Assessment Report of the intergovernmental panel on Climate Change. IPCC, Geneva, Switzerland 151*.

Jacobs, B. W., V. R. Singhal, and R. Subramanian (2010). An empirical investigation of environmental performance and the market value of the firm. *Journal of Operations Management 28*(5), 430–441.

Jaffe, A. B., R. G. Newell, and R. N. Stavins (2005). A tale of two market failures: Technology and environmental policy. *Ecological economics 54*(2-3), 164–174.

Johnson, S. G., K. Schnatterly, and A. D. Hill (2013). Board composition beyond independence: Social capital, human capital, and demographics. *Journal of Management 39*(1), 232–262.

Johnstone, N., I. Haščič, and D. Popp (2010). Renewable energy policies and technological innovation: evidence based on patent counts. *Environmental and resource economics 45*(1), 133–155.

Karpoff, J. M., J. R. Lott, Jr, and E. W. Wehrly (2005). The reputational penalties for environmental violations: Empirical evidence. *The Journal of Law and Economics 48*(2), 653–675.

Kharif, O. (2016). No credit history? no problem. lenders are looking at your phone data. *Bloomberg. com*.

Kim, J.-B., Z. Wang, and L. Zhang (2016). Ceo overconfidence and stock price crash risk. *Contemporary Accounting Research 33*(4), 1720–1749.

Kirkman, B. L., K. B. Lowe, and C. B. Gibson (2006). A quarter century of culture's consequences: A review of empirical research incorporating hofstede's cultural values framework. *Journal of International Business Studies 37*(3), 285–320.

Kirkman, B. L., K. B. Lowe, and C. B. Gibson (2017). A retrospective on culture's consequences: The 35-year journey. *Journal of International Business Studies 48*(1), 12–29.

Klassen, R. D. and C. P. McLaughlin (1996). The impact of environmental management on firm performance. *Management science 42*(8), 1199–1214.

Koshevoy, G. and K. Mosler (1996). The lorenz zonoid of a multivariate distribution. *Journal of the American Statistical Association 91*(434), 873–882.

Kreiser, P. M., L. D. Marino, P. Dickson, and K. M. Weaver (2010). Cultural influences on entrepreneurial orientation: The impact of national culture on risk taking and proactiveness in smes. *Entrepreneurship theory and practice 34*(5), 959–984.

Kuhn, M., K. Johnson, et al. (2013). *Applied predictive modeling*, Volume 26. Springer.

Lee, N. T. (2018). Detecting racial bias in algorithms and machine learning. *Journal of Information, Communication and Ethics in Society*.

Leung, K. and M. W. Morris (2015). Values, schemas, and norms in the culture–behavior nexus: A situated dynamics framework. *Journal of International Business Studies 46*(9), 1028–1050.

Lev, B., B. Sarath, and T. Sougiannis (2005). R&D reporting biases and their consequences. *Contemporary Accounting Research 22*(4), 977–1026.

Lev, B. and T. Sougiannis (1996). The capitalization, amortization, and value-relevance of R&D. *Journal of accounting and economics 21*(1), 107–138.

Li, K., D. Griffin, H. Yue, and L. Zhao (2013). How does culture influence corporate risk-taking? *Journal of corporate finance 23*, 1–22.

Lievenbrück, M. and T. Schmid (2014). Why do firms (not) hedge?—novel evidence on cultural influence. *Journal of Corporate Finance 25*, 92–106.

Liu, X. (2016). Corruption culture and corporate misconduct. *Journal of Financial Economics 122*(2), 307–327.

Markus, H. R. and S. Kitayama (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological review 98*(2), 224.

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics 160*(4), 835–850.

Martin, K. D., J. B. Cullen, J. L. Johnson, and K. P. Parboteeah (2007). Deciding to bribe: A cross-level analysis of firm and home country influences on bribery activity. *Academy of Management Journal 50*(6), 1401–1422.

Michalos, A. C. and P. M. Hatch (2019). Good societies, financial inequality and secrecy, and a good life: from aristotle to piketty. *Applied Research in Quality of Life*, 1–50.

Molnar, C. (2020). *Interpretable machine learning*. Lulu. com.

Mourouzidou-Damtsa, S., A. Milidonis, and K. Stathopoulos (2019). National culture and bank risk-taking. *Journal of Financial Stability 40*, 132–143.

Munnell, A. H., G. M. Tootell, L. E. Browne, and J. McEneaney (1996). Mortgage lending in boston: Interpreting hmda data. *The American Economic Review*, 25–53.

Narver, J. C. (1971). Rational management responses to external effects. *Academy of Management Journal 14*(1), 99–115.

Newell, R. G., A. B. Jaffe, and R. N. Stavins (1999). The induced innovation hypothesis and energy-saving technological change. *The Quarterly Journal of Economics 114*(3), 941–975.

Noailly, J. and R. Smeets (2015). Directing technical change from fossil-fuel to renewable energy innovation: An application using firm-level patent data. *Journal of Environmental Economics and Management 72*, 15–37.

Nyborg, K. and T. Zhang (2013). Is corporate social responsibility associated with lower wages? *Environmental and Resource Economics 55*(1), 107–117.

Ohlson, J. A. (1989). Accounting earnings, book value, and dividends: The theory of the clean surplus equation. *Unpublished paper*.

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary accounting research 11*(2), 661–687.

Oriani, R. and M. Sobrero (2008). Uncertainty and the market valuation of R&D within a real options logic. *Strategic Management Journal 29*(4), 343–361.

OSTP (2016). Of artificial intelligence. In *Robotics, Privacy and Data Protection: Room document for the 38th International Conference of Data Protection and Privacy Commissioners*.

O'Neil, C. Weapons of math destruction.

Pakes, A. (1985). On patents, R&D, and the stock market rate of return. *Journal of political economy 93*(2), 390–409.

Palmer, K., W. E. Oates, and P. R. Portney (1995). Tightening environmental standards: the benefit-cost or the no-cost paradigm? *Journal of economic perspectives 9*(4), 119–132.

Pandit, S., C. E. Wasley, and T. Zach (2011). The effect of research and development (R&D) inputs and outputs on the relation between the uncertainty of future operating performance and R&D expenditures. *Journal of Accounting, Auditing & Finance 26*(1), 121–144.

Parsons, C. A., J. Sulaeman, and S. Titman (2018). The geography of financial misconduct. *The Journal of Finance 73*(5), 2087–2137.

Peterson, M. F. and T. S. Barreto (2018). Interpreting societal culture value dimensions. *Journal of International Business Studies 49*(9), 1190–1207.

Pfeffer, J. and C. T. Fong (2005). Building organization theory from first principles: The self-enhancement motive and understanding power and influence. *Organization Science 16*(4), 372–388.

Popp, D. (2002). Induced innovation and energy prices. *American economic review 92*(1), 160–180.

Porter, M. and C. Van der Linde (1995). Green and competitive: Ending the stalemate. *The Dynamics of the eco-efficient economy: environmental regulation and competitive advantage 33*.

Porter, M. E. (1992). Capital disadvantage: America's failing capital investment system. *Harvard business review 70*(5), 65–82.

Puspitasari, E., C. Sukmadilaga, H. Suciati, R. F. Bahar, and E. K. Ghani. The effect of financial secrecy and ifrs adoption on earnings quality: A comparative study between indonesia, malaysia and singapore.

Reinhardt, F. L. (1999). Bringing the environment down to earth. *Harvard business review 77*(4), 149–57.

Richardson, G. (2008). The relationship between culture and tax evasion across countries: Additional evidence and extensions. *Journal of International Accounting, Auditing and Taxation 17*(2), 67–78.

Russo, M. V. and P. A. Fouts (1997). A resource-based perspective on corporate environmental performance and profitability. *Academy of management Journal 40*(3), 534–559.

Schuler, R. S. and N. Rogovsky (1998). Understanding compensation practice variations across firms: The impact of national culture. *Journal of International business studies 29*(1), 159–177.

Seele, P., C. Dierksmeier, R. Hofstetter, and M. D. Schultz (2019). Mapping the ethicality of algorithmic pricing: A review of dynamic and personalized pricing. *Journal of Business Ethics*, 1–23.

Shane, P. B. and B. H. Spicer (1983). Market response to environmental information produced outside the firm. *Accounting Review*, 521–538.

Shao, L., C. C. Kwok, and R. Zhang (2013). National culture and corporate investment. *Journal of International Business Studies 44*(7), 745–763.

Shapley, L. (1953a). Quota solutions op n-person games1. *Edited by Emil Artin and Marston Morse*, 343.

Shapley, L. (1953b). A value for n-person games, volume ii, chapter contributions to the theory of games.

Shupp, R. S. and A. W. Williams (2008). Risk preference differentials of small groups and individuals. *The Economic Journal 118*(525), 258–283.

Sorley, W. R. (1885). *On the Ethics of Naturalism*. W. Blackwood and Sons.

Sougiannis, T. (1994). The accounting based valuation of corporate r&d. *Accounting review*, 44–68.

Spicer, B. H. (1978). Investors, corporate social performance and information disclosure: An empirical study. *Accounting Review*, 94–111.

Štrumbelj, E. and I. Kononenko (2010). An efficient explanation of individual classifications using game theory. *The Journal of Machine Learning Research 11*, 1–18.

Štrumbelj, E. and I. Kononenko (2014). Explaining prediction models and individual predictions with feature contributions. *Knowledge and information systems 41*(3), 647–665.

Tanzi, V. (1980). Inflationary expectations, economic activity, taxes, and interest rates. *The American Economic Review 70*(1), 12–21.

Tarashev, N., K. Tsatsaronis, and C. Borio (2016). Risk attribution using the shapley value: Methodology and policy applications. *Review of Finance 20*(3), 1189–1213.

Tobin, J. (1969). A general equilibrium approach to monetary theory. *Journal of money, credit and banking 1*(1), 15–29.

Tsakumis, G. T., A. P. Curatola, and T. M. Porcano (2007). The relation between national cultural dimensions and tax evasion. *Journal of international accounting, auditing and taxation 16*(2), 131–147.

Tung, R. L. and G. K. Stahl (2018). The tortuous evolution of the role of culture in ib research: What we know, what we don't know, and where we are headed. *Journal of International Business Studies 49*(9), 1167–1189.

Turban, D. B. and D. W. Greening (1997). Corporate social performance and organizational attractiveness to prospective employees. *Academy of management journal 40*(3), 658–672.

UN, U. N. (2020). Tax abuse, money laundering and corruption plague global finance. *Policy Brief July 2020*.

USACM (2017). Statement on algorithmic transparency and accountability. *Association for Computing Machinery US Public Policy Council*.

Van den Steen, E. (2004). Rational overoptimism (and other biases). *American Economic Review 94*(4), 1141–1151.

Villani, C. (2018). For a meaningful artificial intelligence: towards a french and european strategy.

Vinas, F. (2021). How financial shocks transmit to the real economy? banking business models and firm size. *Journal of Banking & Finance 123*, 106009.

Vitell, S. J., S. L. Nwachukwu, and J. H. Barnes (1993). The effects of culture on ethical decision-making: An application of hofstede's typology. *Journal of business Ethics 12*(10), 753–760.

Volkema, R. J. (2004). Demographic, cultural, and economic predictors of perceived ethicality of negotiation behavior: A nine-country analysis. *Journal of Business Research 57*(1), 69–78.

Von Arx, U. and A. Ziegler (2014). The effect of corporate social responsibility on stock performance: New evidence for the USA and Europe. *Quantitative Finance 14*(6), 977–991.

Walley, N. and B. Whitehead (1994). It's not easy being green. *Reader in Business and the Environment 36*, 81.

Watts, L. L., L. M. Steele, and D. N. Den Hartog (2020). Uncertainty avoidance moderates the relationship between transformational leadership and innovation: A meta-analysis. *Journal of International Business Studies 51*(1), 138–145.

Westjohn, S. A., P. Magnusson, Y. Peng, and H. Jung (2021). Acting on anger: Cultural value moderators of the effects of consumer animosity. *Journal of International Business Studies 52*(8), 1591–1615.

Wexler, R. (2017). How companies hide software flaws that impact who goes to prison and who gets out. *Washington Monthly*.

Williams, K. M., C. Nathanson, and D. L. Paulhus (2010). Identifying and profiling scholastic cheaters: Their personality, cognitive ability, and motivation. *Journal of Experimental Psychology: Applied 16*(3), 293.

Zhang, D. H. and P. S. Willen (2021). Do lenders still discriminate? a robust approach for assessing differences in menus. Technical report, National Bureau of Economic Research.

Ziegler, A., M. Schröder, and K. Rennings (2007). The effect of environmental and social performance on the stock performance of European corporations. *Environmental and Resource Economics 37*(4), 661–680.