APPLIED MICROECONOMIC ESSAYS ON: IMMIGRATION, LABOR AND WELFARE

A Dissertation

by

YINJUNJIE ZHANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Marco A. Palma
Committee Members,	Ximing Wu
	Yu Zhang
	Yonghong An
Head of Department,	C. Parr Rosson III

August 2018

Major Subject: Agricultural Economics

ABSTRACT

This dissertation encompasses three different topics in applied microeconomics with the common thread of understanding the impact of public policy on economic behavior, choice, and welfare. By employing causal inference methods, I identify the intended and unintended effects of immigration policies on crime rates and labor market preferences. By applying a nonparametric identification technique, I evaluate the measurement error in reported subjective well-being, an alternative criterion gaining the attention of more researchers to evaluate policy interventions.

In the first essay, I study the effect of Alabama's HB56 immigration law on crime. Alabama HB 56 passed in 2011, is considered to be the strictest anti-illegal immigration bill in the United States. By using the synthetic control method to create a counter-factual Alabama, this chapter provides suggestive evidence of heterogeneous causal effects of Alabama HB 56 on crime. Compared to the synthetic group, the rate of violent crime increased as a response to Alabama HB 56, while there was no significant change in property crime rate after the law was implemented. A placebo test is performed to demonstrate the robustness of the results.

In the second essay, I investigate the impact of the H-1B cap exemption on Ph.D. labor markets. The American Competitiveness in the Twenty-first Century Act of 2000 (AC 21) eliminated the H-1B cap for foreign employees of academic, non-profit and government research organizations. This act potentially affects the job preferences of newly graduated foreign Ph.D students. Choosing a career in an uncapped H-1B qualified entity makes the foreign-born Ph.D. graduates to circumvent the risk of facing the fiercely competitive H-1B application process and possibly avoiding potential losses due to a visa rejection. I use data from the Census of Ph.D. graduates to examine the policy impact on academic and industry labor markets in the United States. The results show that Ph.D. graduates with temporary visas are 5% more likely to pursue a job in academia, and 3-4% less likely to choose a job in industry. Placebo and falsification tests on post-doctoral participation further exclude other external changes that could possibly affect the job market.

As a measure of welfare, the growth of subjective well-being (SWB) is among the most critical aims in human society. However, self-reported well-being is potentially subject to significant misclassification errors. In the final essay, I employ a recently developed method from the measurement error literature to correct measures of reported happiness in 80 countries. I find that misclassification errors are correlated with prevalent religious beliefs and countries' economic development stages, along with other individual characteristics. By utilizing the corrected SWB, I further reexamine the Easterlin paradox and modified-Easterlin hypothesis. The findings indicate that although *reported* SWB is not associated with GDP *per capita*, the *corrected* measure of SWB is. I find no evidence for a happiness satiation point as defined in previous studies.

DEDICATION

To my parents, Zhiquan and Mimi.

ACKNOWLEDGMENTS

I am deeply grateful to my advisor Dr. Marco A. Palma for his professional guidance, warmest support and encouragement. The research environment he has created assists me in many aspects. I would have never become as productive and passionate about my research without his supervision and high level of expectations.

I am indebted to Dr. Yonghong An for his suggestions and advice, especially in my third essay. I thank Dr. Ximing Wu and Dr. Yu Zhang for their valuable comments on my research. During the journey of pursuing my Ph.D., I have received countless advices and insights from friends and coauthors within our team and professors outside of the agricultural economics department. I acknowledge them sincerely.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Marco A. Palma (Advisor) and Professors Ximing Wu and Yu Zhang of the Agricultural Economics Department and Professor Yonghong An of the Economics Department.

Chapter 1 was conducted with data from the American Community Survey (ACS), Bureau of Labor Statistics (BLS), Current Population Survey (CPS), FBI's Uniform Crime Reporting (UCR) statistics, Gallups, and Substance Abuse and Mental Health Services Administration. The restricted data analyzed in Chapter 2 was provided by the National Science Foundation National Center for Science and Engineering Statistics (NSF/NCSES). The Data analyzed in Chapter 3 was obtained from the European Values Surveys (EVS), World Value Survey Database (WVS), Penn World Table and The World Factbook.

The analyses depicted in Chapters 1 and 3 were conducted in part by Zhicheng Xu of the Economic Department, Henan University. All other work for this dissertation was independently completed by the student.

Funding Sources

My graduate studies were supported by a fellowship from Texas A&M University.

NOMENCLATURE

ACS	American Community Survey
BLS	Bureau of Labor Statistics
CPS	Current Population Survey
EVS	European Values Surveys
FCC	Freedom of Choice and Control in Life
LS	Life Satisfaction
NSF/NCSES	National Science Foundation National Center for Science and Engineering Statistics
SWB	Subjective Well-being
UCR	FBI's Uniform Crime Reporting
WVS	World Value Survey

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

Evaluation of public economic policies has historical roots in social and political philosophy. From Aristotle's Politics to John Stuart Mill's On Liberty and Utilitarianism, ancient philosophers developed ideas on the government's role in improving society and whether a government should interfere with individual's decision-making process. In modern society, difficulties in achieving social reconciliation on matters of public concern are also mainly due to the competing philosophical principles held by social groups. The problems can be of economic, social and political nature. Policy instruments serve to mitigate differences among social classes on these issues.

As a continuous process, policy making generates several feedback loops. Evaluating the economic and social impact of policy intervention plays a critical role in guaranteeing the implementation of goals set by program designers and policy makers. In this sense, the concept of program evaluation not only refers to the results of policy interventions, but more broadly to the efficiency and efficacy of the governmental decision-making process.

While the ultimate goal of policy interventions is to improve social well-being, the criterion for improvement has been extended from objective to subjective measures in economic studies. With limited emphasis on subjective measures, in the past policy makers neglected the possibility that policy interventions, e.g. welfare programs, could negatively affect an individual's subjective well-being (SWB).[1] As a result, many individuals eligible for a program refuse to participate when only objective indicators are considered to define well-being improvement.[2]

This doctoral research is an attempt to shed light on two aspects of public policy analysis. Using causal inference techniques, the first two studies focus on identifying the socio-economic impacts of immigration laws of two types: one is associated with undocumented immigrants, while the other one is linked to foreign-born high-skilled labor market participants. In the third study, with a newly developed eigenvalue-eigenvector decomposition technique, I identify the prevalent

[1]Recipients may derive disutility from feelings of lack of self-respect and negative self-characterizations of participating in a welfare program (Horan and Austin, 1974).

[2]This phenomenon, introduced by Moffitt (1983) in his seminal paper, is known as "Welfare Stigma".

measurement errors in self-reported subjective well-being, one critical issue when using SWB as an indication of welfare improvement in policy evaluation.

Viewed as one focus of public concern, immigration is at the center of political debate. Identifying the intended and unintended impacts of immigration policies on different aspects of society has significant repercussions. In the first essay, I identify an anti-illegal immigrant bill's initial effect on crime rate. In 2011, Alabama enacted the Beason-Hammon Alabama Taxpayer and Citizen Protection Act (Alabama HB 56), which was considered the nation's strictest anti-illegal immigration bill (Fausset, 2011). The bill imposes extreme restrictions on undocumented immigrants in Alabama and set limits on every aspect of their lives, including enrollment in schools and job market participation. Our study is motivated by a recent interest in the literature to evaluate anti-illegal immigrant laws (Bohn et al., 2014; Hoekstra and Orozco-Aleman, 2017), and its impact on crime rate (Bell et al., 2013). To evaluate the impact of Alabama HB 56 on crime this paper uses the synthetic control approach, which was first employed by Abadie and Gardeazabal (2003). I create a counter-factual synthetic Alabama based on crime reports, economic indicators, and demographic characteristic during the period 1998-2014. Variables used in constructing the synthetic Alabama are chosen following the literature on crime rate (Blau and Blau, 1982; Bailey, 1984; Fajnzylber et al., 2002). Particularly, to proxy the risk of being arrested and punished, I use police presence and capital sentence legality accessed from the Bureau of Labor Statistics. I used the admission cases of primary substance use of marijuana and alcohol, accessed from the Substance Abuse and Mental Health Services Administration, as a proxy for illicit drug use. Gallup's polling data for practicing religion and church attendance are employed to account for cultural characteristics. State level economic indicators such as Gini index and unemployment status are obtained in the American Community Survey and Current Population Survey. Estimation results provide suggestive evidence of heterogeneous causal effects of Alabama HB 56 on crime. Compared with the synthetic counterpart, Alabama HB 56 contributed to an increase in violent crime rates, while there was no significant change in property crime rates after the act. The results can be linked to the seminal work of Gary Becker (Becker, 1968), who modeled criminal behavior from a rational decision analysis based on the benefits and opportunity costs of crime. In general, unemployment status decreases the opportunity cost of criminal behavior. Fleisher (1966) and Ehrlich (1973) document a significant causal effect of unemployment on criminal activities in the US. Since the Alabama HB restricts undocumented immigrants from taking job positions, it is expected to increase criminal activities. Another possible mechanism is associated to the growing literature on expressive value of law, which suggests the act of enacting a particular law serves the function of shaping norms or prescribed attitude towards behavior (Bursztyn and Jensen, 2017). The adoption of an anti-immigration law will send out a signal of increased tolerance to discrimination behavior against undocumented workers, which can also potentially generate tensions or conflicts that may result in violent crimes against undocumented immigrants. Whether the increase in violent crimes is committed by undocumented immigrants or against them, and through which underlying mechanism merit further investigation using richer data.

In the second essay, I leverage the H-1B policy shock generated by the AC 21 as an exogenous variation and examine the subsequent changes in foreign Ph.D graduates' job market preferences. The American Competitiveness in the Twenty-first Century Act (AC 21) was signed into law in October, 2000 (U.S. Congress, 2000). It eliminates the cap on the number of H-1B visas granted for applicants in universities, non-profit and government research institutions. As a result, choosing a career in an uncapped H-1B qualified entity means to circumvent the risk of facing the fiercely competitive H-1B application process and possibly avoiding potential losses due to a visa rejection. Accordingly, AC 21 would impose a direct effect on foreign-born Ph.D. graduates' job preference in academia, while having an indirect (or second degree) effect on the job preference in industry. Our study is motivated by a profound linkage between the H-1B program and high-skilled labor markets in the United States. One key issue surrounding the immigration debate is whether high-skilled immigrants have a complementary or displacing effect on domestic workers. Most immigrant scientists and engineers started their careers in the United States via this visa program and they are now making substantial contributions to the economy and advancing technological development (Peri, 2007; Anderson and Platzer, 2006; Wadhwa et al., 2007). In this sense, a relaxed visa policy for high-skilled immigrants allows experts from a wide range of fields to stay in the US labor market after completing their studies, producing benefits over the long run. Yet, the debate about the short-run effects of the H-1B visa program is still unsettled. A crowding-out influence from the influx of high-skilled immigrants on domestic workers is supported by the literature (Borjas, 2005; Islam, 2009; Ottaviano and Peri, 2012; Borjas and Doran, 2012). A crucial challenge for detecting this effect relies on overcoming the endogeneity of the behavior of immigrants in the job market. An evaluation of how immigration policy changes influence job market preferences of high-skilled immigrants paves the way to accurately identify the impact of foreign workers on domestic workers. To quantitatively identify the treatment effect on job market choices, we establish our estimation model using individual level variation in visa status and graduation time, before and after the policy shock. Due to idiosyncratic features of each doctorate academic field, our baseline specification includes both major and year fixed effects. In order to allow for different disciplines to have distinct trajectories over time, we estimate models that incorporate major-specific linear trends. Specifications extending to incorporate field-by-year fixed effects to allow majors of different fields to accommodate differential shocks are also investigated. Our Difference-in-Difference estimates suggest a strong causal relationship between AC 21 and the job preferences of Ph.D holders in academia and industry labor markets. We find that after the implementation of AC 21, foreign Ph.D graduates are 5% more likely to pursue a job in academia, and 3-4% less likely to choose an industry job compared to U.S. domestic graduates. Placebo experiments and falsification tests exclude the possibility of other external changes around the same time period driving these results.

In the third essay, we examine the potential misclassification errors in self-reported happiness data and explore the characteristics of countries and demographic groups with substantial misclassification errors. The well-known Easterlin paradox and modified Easterlin hypothesis are re-examined with the corrected SWB data. The pursuit of happiness has been paramount for humankind throughout history. It has evoked the attention of philosophers from Epicurus and Aristotle to Jeremy Bentham and John Stuart Mill. Evaluating social well-being is a central topic in economics. A recent and growing literature argues that income and consumption are not ideal measures of welfare (see Fleurbaey (2009) for a review). Subjective well-being (SWB) can arguably be used as a standard for measuring social welfare and progress. Recently, it is gaining attention in economic research on policy evaluation. The most frequently used method to elicit SWB is to directly ask people about their feelings of happiness and overall life satisfaction (e.g., Frey and Stutzer, 2000, 2002, 2010; Diener, 2000). However, compared to objective indicators of social welfare, SWB is subject to substantial misreporting or measurement errors (e.g., Hagedorn, 1996; OECD, 2013; Diener et al., 2013; Sheridan et al., 2015; Chetty, 2015). Although the ample evidence of measurement errors is documented in self-reported happiness, this issue has been mostly ignored by empirical practitioners. To address this concern, we apply a novel closed-form identification and estimation method first proposed by Hu (2008) and developed in Feng and Hu (2013). In general, the methodology employs an eigenvalue-eigenvector decomposition technique to establish a closed-form identification for the misclassification matrix of SWB, which is constructed a priori using the conditional distribution of reported happiness. By pre-multiplying the inverse of the identified misclassification matrix by the unconditional distribution of reported SWB, the true latent SWB distribution is directly estimated. The identification strategy in our article is implemented closely in the spirit of Feng and Hu (2013). For computation time reasons, we select 80 countries based on geographical diversity and consistency of survey data across periods in the Integrated European/World Value Survey (EVS/WVS). We use reported "feeling of happiness" in EVS/WVS as the direct measurement of SWB, while "life satisfaction" (LS) and "freedom of choice and control in life" (FCC) serve as two repeated measures of SWB. Under relatively weak assumptions, we are able to obtain the latent SWB probability distributions by country and demographic groups. Another motivation of our study is related to the well-known "Easterlin paradox" (Easterlin, 1974, 1995; Blanchflower and Oswald, 2004). *The Easterlin paradox* posits that although there is a positive correlation between individual income and measures of SWB, aggregate happiness is not significantly associated with GDP *per capita*. Meanwhile, a modified version of the Easterlin's hypothesis has also been proposed. It argues that the correlation between income and SWB only exists for individuals with income below a certain threshold, beyond which income is no longer related to happiness (Frey and Stutzer, 2002; Clark et al., 2008; Di Tella and MacCulloch, 2008; Di Tella et al., 2010). Our analysis produces the following major findings. We find that self-reported happiness elicited from survey questions has substantial misclassification errors. Religious belief and the development stage of a country play critical roles in determining the magnitude of misclassification errors in reported SWB. We revisit the Easterlin paradox and find that based on a country level analysis there is no evidence supporting neither the original hypothesis nor the modified version of the Easterlin's paradox when using the *corrected* measure of happiness, although the *reported* (uncorrected) SWB provides support to the original Easterlin paradox. These findings imply that, to use SWB measures in evaluating economic development and welfare, they have to be corrected for measurement errors.

2. UNINTENDED EFFECTS OF THE ALABAMA HB 56 IMMIGRATION LAW ON CRIME: A PRELIMINARY ANALYSIS*

2.1 Introduction

Over the last decade, illegal immigration has received considerable attention from state governments. In 2007, Arizona passed the Legal Arizona Workers Act (LAWA), which was followed in 2010 by SB 1070; the harshest U.S. act against undocumented immigrants at the time. One year later Alabama enacted the Beason-Hammon Alabama Taxpayer and Citizen Protection Act (Alabama HB 56), which is now considered the nation's strictest anti-illegal immigration bill (Fausset, 2011). Both legislative acts were enacted to restrict the enrollment of undocumented immigrants in schooling and in the job market. The consequences of these anti-illegal immigration acts have received much attention from governments and researchers. Despite the growing interest in the relationship between illegal immigration and crime environment, the evaluation of the causal effect of those policies is subject to debate. While evidence of positive relationship between immigration and crime has been found in some recent studies, such as Spenkuch (2013), a major vast literature provides suggestive evidence that immigrant legalization has contributed to the decline of crime (Baker, 2015; Mastrobuoni and Pinotti, 2015; Pinotti, 2017). Meanwhile, some previous studies suggest that immigration has no effect on crime (Butcher and Piehl, 1998; Chalfin, 2013).

This paper evaluates the impact of Alabama HB 56 on crime using the synthetic control approach, which was first employed by Abadie and Gardeazabal (2003) to estimate the consequences of the conflict in Basque, Spain. They constructed a synthetic 'Basque' using a weighted average of other provinces in Spain according to similarities in economic and demographic indicators. Following their approach, we create a counterfactual synthetic Alabama based on crime reports, economic indicators, and demographic characteristic during the period 1998-2014. The results provide suggestive evidence of heterogeneous causal effects of Alabama HB 56 on crime. Com-

*This chapter is reprinted from *Economics Letters*, 147, Zhang, Yinjunjie, Marco A. Palma, and Zhicheng Phil Xu, "Unintended effects of the Alabama HB 56 immigration law on crime: A preliminary analysis.", 68-71, Copyright (2016), with permission from Elsevier.

pared with the synthetic counterpart, Alabama HB 56 contributed to an increase in violent crime rates, while there was no significant change in property crime rates after the act.

Our study is motivated by a recent interest in the literature to evaluate anti-illegal immigrant laws. For instance, using a synthetic control method, Bohn et al. (2014) find that the 2007 Legal Arizona Workers Act substantially reduced the proportion of undocumented immigrants in the population of Arizona. Hoekstra and Orozco-Aleman (2017) examined the effect of Arizona SB 1070 on the undocumented immigrants' individual decisions regarding their migration destination. Their results show that the passage of the bill significantly reduced unauthorized immigration to Arizona by 70%. Our study focuses on the causal effect of Alabama HB 56 on crime, rather than the direct effect on the proportion of unlawfully present immigrants.

Bell et al. (2013) conducted a similar study where they examined the relationship between crime and immigration in the UK during the 1990s and 2000s. They found that preventing asylum seekers from finding jobs resulted in increases in property crime, but had no impact on violent crime. We use a different estimation framework and focus on the evaluation of the treatment effect of anti-illegal immigrant laws on crime in Alabama. In contrast to their finding, Alabama HB 56 increased violent crime but had no impact on property crime.

Our results can be linked to literature after Becker (1968), who modeled criminal behavior from a rational decision analysis based on the benefits and opportunity costs. In general, unemployment status decreases the opportunity cost of criminal behavior. Fleisher (1966) and Ehrlich (1973) documented the significant causal effect of unemployment on criminal activities in the US. Since the Alabama HB 56 restricts undocumented immigrants from taking job positions, it is expected to increase criminal activities. However, there is also literature suggesting that increasing police force can reduce crime (Di Tella and Schargrodsky, 2004; Levitt, 2002). According to this view, the increase in immigration police force due to Alabama HB 56 may reduce crime. In spite of the contradictory predictions, our results support the former; the anti-immigrant law is more likely to increase violent crime.

The rest of this paper proceeds as follow. Section 2.2 introduces the institutional background.

Section 2.3 describes the methodology and data. Section 2.4 presents the results. Section 2.5 performs placebo tests to provide statistical inference, and section 2.6 concludes.

2.2 Institutional Background

The Alabama HB 56 Bill passed in June 2011 imposes extreme restrictions to undocumented immigrants in Alabama. It requires every public elementary and secondary school to determine whether students were born outside of the US or if their parents are undocumented. The Bill makes it a felony for an undocumented immigrant to "enter into any business transaction with a government agency". It also prohibits signing rental agreements or providing housing accommodations for undocumented immigrants. Similar to other immigration acts, Alabama HB 56 also requires "every business entity or employer in the state to enroll in E-Verify", the federal government's online database used to check the employment eligibility of workers. As a consequence, undocumented immigrants have fewer opportunities for education, employment and businesses, thereby lowering opportunity cost for criminal behaviors.

2.3 Methodology and Data

To evaluate the effect of HB 56, one needs to find a comparable counterfactual for Alabama. This paper employs a data-driven method to construct a synthetic Alabama, which is a weighted average over all the control states. In the framework of the synthetic control method, J is the number of states, $j = 1$ is the treated state, Alabama. The rest of the states from $j = 2, ., J$ are the potential control alternatives constituting a "donor pool" to construct a synthetic Alabama. Define Y_{jt}^{C} and Y_{jt}^{T} as the outcome of unit j at time t in the control group and treatment group respectively. Let $W = w_1, w_2, , w_{J-1}$ be the $(J-1) \times 1$ vector of weights, with $0 \leq w_j < 1$ and $w_1 + w_2 + \ + w_{J-1} = 1$. By choosing the proper W^*, the following minimization problem can be solved at $t = T_0$, the pre-intervention period.

$$W^* = argmin_W (Y_{1t}^{T} - Y_{t}^{C} W)'(Y_{1t}^{T} - Y_{t}^{C} W)$$

Then, the treatment effect can be calculated by applying W^* to $t = T_1$ post-intervention period.

Our outcome variables are crime rates at the state level which are obtained from the FBI's Uniform Crime Reporting (UCR) statistics from 1998 to 2014. To construct the synthetic Alabama, we combine state-level police labor force data from 1998 to 2014 from the Bureau of Labor Statistics (BLS) and calculate *per capita* statistics. We further control for state differences in death penalty legality.

Since Fajnzylber et al. (2002) find that cultural characteristics such as religion views and illicit drug use can affect an individual's propensity to crime, we use the percentage of respondents who identify as Christian from Gallup's polling data (2006-2014).

Due to data limitations, we use the number of hospital admission for primary substance abuse as control for illicit drug use. Specifically, we control alcohol and marijuana use per 100,000 habitants (accessed from the Substance Abuse and Mental Health Services Administration).

There is substantial evidence documenting how income inequality contributes to crime behavior, in particular violent crimes (Blau and Blau, 1982). Thus, the Gini index in the American Community Survey (2006-2014), and unemployment status in the Current Population Survey (1998-2014) are also included.

Regarding demographic aspects, we use the proportion of non-citizen Hispanics over 15-year-old without high school education as the proportion of the population most likely to be undocumented (Bohn et al., 2014).

Over the past decade, several states have launched their own immigration regulations. States impacted by similar interventions should be excluded from the donor pool (Abadie et al., 2015), thus we remove all the states that passed E-verify or Omnibus Immigration Legislation (OIL) around the same period.

2.4 Results

2.4.1 Violent crime

While evaluating the impact of Alabama HB56 on violent crime, synthetic Alabama was constructed by a combination of Kentucky (42.1%), Oklahoma (40.6%), Florida (9.6%), and New

Mexico (7.6%), with W^* displayed in the parentheses. Table 2.1 reports the pre-intervention comparison for Alabama and synthetic Alabama. The synthetic Alabama provides a close reproduction of Alabama.

Table 2.1: Crime Rate Predictors: Mean Trends (1998-2010)

	Alabama	**Synthetic**		**Alabama**	**Synthetic**
Violent crime rate	447.82	447.10	Hispanic noncitizen at least 15		
White	0.69	0.82	with less than high school	0.01	0.01
Age under 18	0.25	0.25	Gini index	0.47	0.46
Age 18-44	0.35	0.35	Primary marijuana admission	139.17	94.47
Age 45-64	0.26	0.26	Primary alcohol admission	181.08	227.38
Less than high school	0.24	0.21	Detectives and criminal investigator	22.69	25.67
High school	0.33	0.33	Police and sheriff's patrol officers	206.04	190.99
Christian	0.87	0.83	Security guards	305.97	282.84
Unemployment rate	0.05	0.05	Death penalty	1.00	0.99

Figure 2.1 shows the trends for violent crime cases per 100,000 people in Alabama and synthetic Alabama. The magnitude of the estimated impact of HB 56 is significant in the post-intervention period. In the pre-intervention periods (1998-2010), the violent crime rate for synthetic Alabama is fairly close to the rate in actual Alabama showing a good model fit.

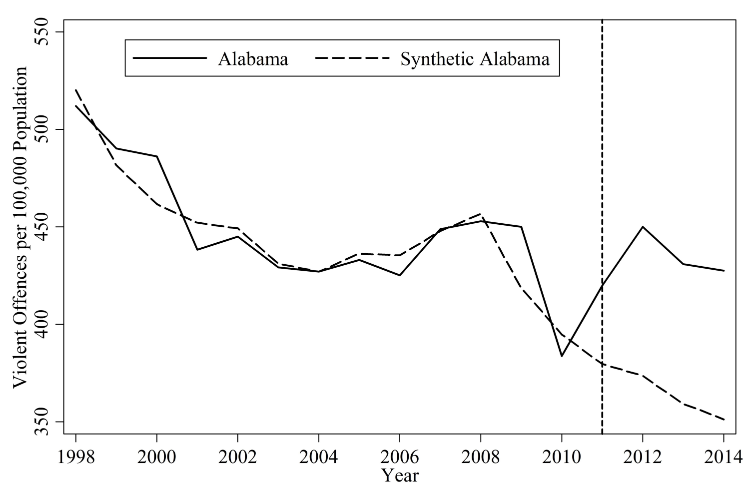


Figure 2.1: Trends in Violent Crime Cases

2.4.2 Property crime

Figure 2.2 displays the total property crime rates from 2002 to 2014. There is no evidence of an impact of the immigration policy. Alabama and synthetic Alabama both experienced a decline in property crime rates during 2008-2014.

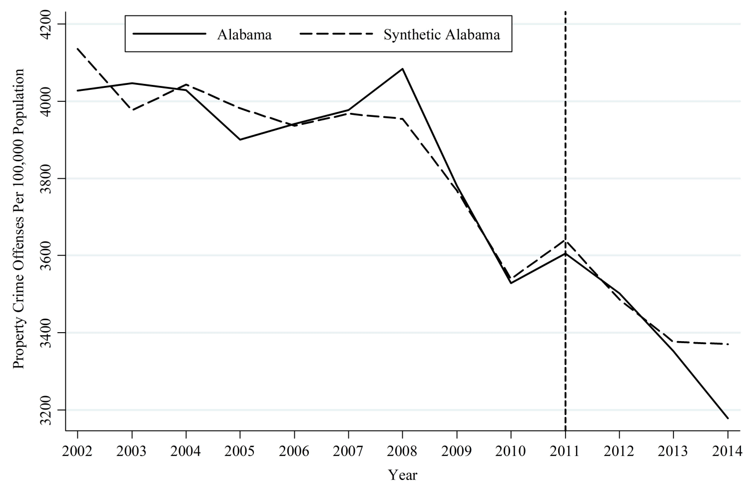


Figure 2.2: Trends in Property Crime Cases

2.5 Placebo Studies and Robustness Test

To validate our finding regarding violent crime, we implement placebo studies and robustness checks for inference (Abadie et al., 2010, 2015). In the placebo studies, we repeatedly assign the intervention to each of the control states in the donor pool which did not enact anti-illegal immigrant laws during the same period. If a significant placebo effect is detected, then the estimated shift for Alabama would be seriously undermined.

Figure 2.3 displays the placebo test excluding states with pre-treatment periods' mean squared prediction error (MSPE) larger than twice of Alabama. MSPE measures the magnitude of the difference in the outcome variable between the treated unit and its synthetic counterpart (Abadie et al., 2010, 2015). The states with large fluctuating MSPE in the pre-treatment period will not provide valid information. The black line denotes the gap in the outcome variable between the treated and synthetic control group, while the gray lines represent the gaps in each of the iterative application of intervention for the units in the donor pool. Twenty-two controls states and Alabama are presented in Figure 2.3. Alabama displays the largest gap-line in the post-intervention period. According to Abadie et al. (2010), the probability of estimating such a magnitude of gap-line under the random treatment assignment is as low as 1/22, which in terms of the inference interpretation means a p-value of lower than 5%.

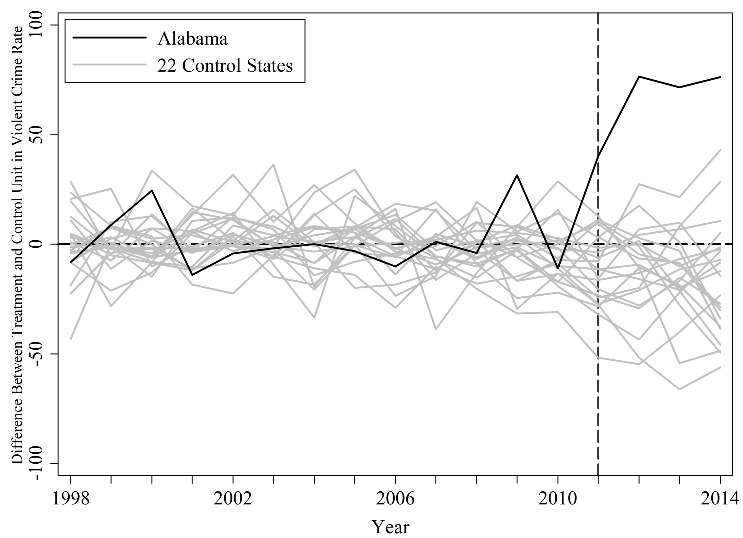


Figure 2.3: Violent Crime Gaps in Alabama and Placebo Tests in 22 Control States

We also perform the placebo tests by examining the distribution of post-pre MSPE ratio across states. A large post-period MSPE does not validate the treatment induced shift unless the post-pre period ratio of MSPE is also large. Figure 2.4 reports how this ratio spreads across Alabama and all the other 37 control states. Alabama has a value of post-period about 27 times larger than the pre-period. If the treatment is randomly assigned to any state in the donor pool, the probability of obtaining a MSPE ratio as large as Alabama's is 1/37 = 0.027.

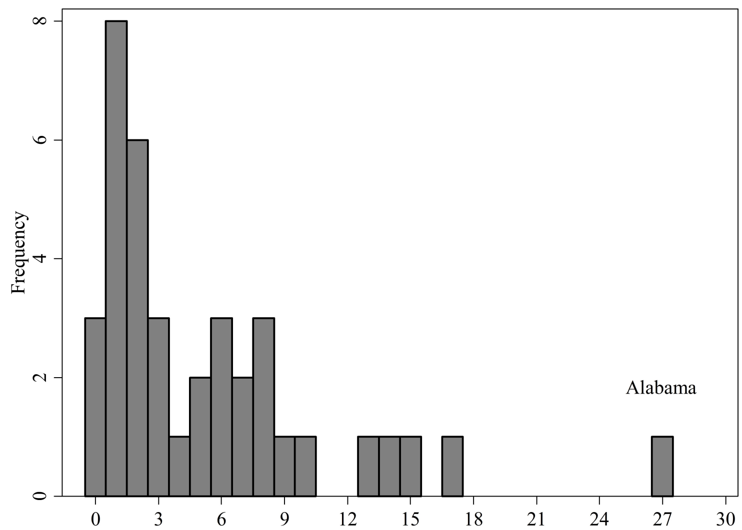


Figure 2.4: Ratio of Post/Pre-HB56 MSPE: Alabama and 37 Control States

We further implement the sensitivity test by leaving out the four states with positive weights one at a time and re-estimating the model. Figure 2.5 plots the four synthetic groups constructed by leaving one state out together with the reproduction of Figure 2.1. As seen, our result for violent crime is robust to the exclusion of any particular state with positive weight. All leave-one-out synthetic groups present similar effects to our findings. By sacrificing some extent of goodness of fit, it demonstrates our result is not driven by any particular state.

2.6 Conclusion

The recent waves of anti-immigrant regulation have been paid interest from researchers and governments. In this paper, we estimate the causal effect of the harshest anti-immigration bill, Alabama HB 56, on crime using a synthetic control approach. We provide suggestive evidence of heterogeneous effects of Alabama HB 56 on violent crime and property crime. Although this anti-immigrant bill did not affect property crime in Alabama, violent crime significantly increased after the bill was enacted.

The intention of this act is to reduce the amount of illegal immigrants living in Alabama and create a better socio-economic environment for the local citizens. However, its consequences

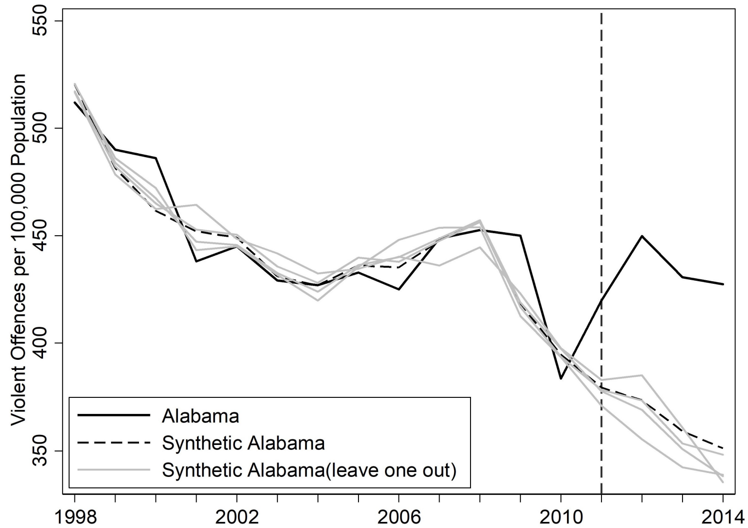


Figure 2.5: Leave-One-Out Distribution of the Synthetic Control for Alabama

are contradictory to the expectations. There is evidence for the negative economic consequences caused by the passage of this act. HB 56 is estimated to cost Alabama as much as $11 billion in economic output and another $264.5 million in tax revenue (Addy, 2012). And in this study we find unintended effect on violent crime rate. This calls for the state government to take the potential negative side effects into consideration while adjusting future immigration-related bills.

Our results can be linked to the literature in which criminal behavior is analyzed based on the benefits and opportunity costs Becker (1968). Since Alabama HB56 restricts the employment of undocumented immigrants, it is expected to increase criminal activities because of lower opportunity cost of criminal behavior, although no effect is found on property crime. On the other hand, increased crimes may also be committed against undocumented immigrants, since undocumented victims have less incentive to report crimes against them. Due to lack of available data, whether the increase in violent crimes is committed by undocumented immigrants or against them merits further investigation.

3. THE IMPACT OF THE H-1B CAP EXEMPTION ON PH.D. LABOR MARKET

3.1 Introduction

Immigration policy is currently one of the most widely discussed topics in government and among regular citizens. It is a sensitive issue that carries strongly polarized views. The majority of immigration arguments are centered around low-skilled, low-paying jobs. We, on the other hand, concentrate on high-skilled labor markets, through which foreign professionals are normally hired using the H-1B visa program. There are competing views among researchers and policy makers on whether the US government should assign more visas to allow foreign high-skilled workers to stay in the US labor market; or make the current immigration policy stricter with the argument that immigrants displace native US workers. Policy changes surrounding the H-1B temporary visa program have been heatedly debated since the program was first implemented in the 1990s.

The H-1B program administrates the eligibility of full-time employment for most US trained high-skilled immigrants. In particular, for newly graduated foreign-born professionals including all academic degree holders, the H-1B visa program is possibly the only path to legally enter into the US labor market.[1] The volume of H-1B petitions is almost always larger than the number of available visas. As a result, the US Citizens and Immigration Services (USCIS) sets quotas for the number of H-1B visas to be issued every year. The cap was 115,000 in 1999 and 2000; 195,000 from 2001 to 2003; and 65,000 after 2004, with an additional 20,000 for applicants holding at least a Master's degree.[2] For some years, the excessive number of petitions even prompted the USCIS to utilize lottery procedures to arrange for visa assignments.

The American Competitiveness in the Twenty-first Century Act (AC 21) was signed into law in October, 2000 (U.S. Congress, 2000). It had one key provision: the H-1B visa would not be subjected to the annual quotas for applicants employed by higher education institutions, nonprofit research organizations, and government research organizations. The enactment of AC 21 marks

[1] We refer here to an employment-based visa. It is also possible for foreign workers to enter the labor market through marriage with a US citizen.

[2] https://www.uscis.gov/

the beginning of the H-1B exemption in the numerical cap for research oriented institutions. As a result, it potentially affects the job preferences of non-citizen college graduates seeking to stay in the United States after graduation. Choosing a career in an uncapped H-1B qualified entity makes the foreign-born Ph.D. graduates to circumvent the risk of facing the fiercely competitive H-1B application process and possibly avoiding potential losses due to a visa rejection. The mechanism to establish a potential causal relationship is further implemented through the clause stating that "any individual who is on an uncapped H-1B visa will be counted towards the cap if he/she switches to a job that is subject to a cap" (U.S. Congress, 2000). According to these terms, AC 21 would impose a direct effect on the job preference in academia, while having only an indirect (or second degree) effect on the job preference in industry.

The intent-to-treat effect from AC 21 possibly affects all students who graduated with any academic degree, but jobs in academia or research institutions in general require a doctoral degree. Thus, we concentrate on the identification of causal effects on Ph.D holders as they participate in high-skilled labor markets and are important drivers of innovation and discovery in science and industry (Cohen et al., 2002; Stuen et al., 2012).

Our Difference-in-Difference estimates suggest a strong causal relationship between AC 21 and the job preferences of Ph.D holders in academia and industry labor markets. We find that after the implementation of AC 21, foreign Ph.D graduates are 5% more likely to pursue a job in academia, and 3-4% less likely to choose an industry job compared to U.S. domestic graduates. Placebo experiments and falsification tests exclude the possibility of other external changes around the same time period driving these results.

The rest of the paper proceeds as follows. In the following section, we document the literature and motivation. Section 3 provides a brief overview of the institutional details of AC 21. Section 4 provides an overview of the data and presents descriptive results of our analysis. Section 5 introduces our identification strategy, where we leverage individual level variation in the visa status to identify the effects of AC 21 on job placement. Section 6 reports the results. In Section 7, we discuss the two inference strategies that we use and present results for the placebo experiments and falsification test. Section 8 explores heterogeneous effects by doctoral field and section 8 concludes.

3.2 Literature and Motivation

The linkage between the H-1B program and high-skilled labor markets is profound. Most immigrant scientists and engineers started their careers in the United States via this visa program and they are now making substantial contributions to the US economy and advancing technological development (Peri, 2007; Anderson and Platzer, 2006; Wadhwa et al., 2007).[3] In this sense, a relaxed visa policy for high-skilled immigrants allows experts from a wide range of fields to stay in the US labor market after completing their studies, producing benefits over the long run.

Yet, the debate about the short-run effects of the H-1B visa program is still unsettled. On one hand, high-skilled immigrants provide both critical contributions in fundamental research and auxiliary skills in management and entrepreneurship. For instance, Kerr and Lincoln (2010) found that a 10% increase in the H-1B population resulted in around 4% higher growth in the number of noncitizen scientists and engineers without displacement impacts on their native counterparts. Further, Hunt and Gauthier-Loiselle (2010) document that an increase in the share of immigrant college graduates induces a sizable positive spill-over effect on innovation. By leveraging the lottery assignment of H-1B as an exogenous intervention in computer-related occupation markets, Peri et al. (2014a) showed that negative shocks in H-1B assignment generated from the lottery process reduced employment and the growth rate of wages for native workers; this result suggests a complementary relationship between native and foreign workers. Similarly, Islam et al. (2017) found that the increase of skilled immigrants in a particular group raised the wages of both native and immigrants from that group. Using historical data, Moser et al. (2014) document a productivity boost on fellow Chemistry researchers following the immigration of Jewish scientists from Nazi Germany.

On the other hand, a crowding-out influence from the influx of high-skilled immigrants on domestic workers is also supported by the literature. This effect was found either through direct

[3] Kerr (2013) provides a review of the contributions of immigrants to US innovation and entrepreneurship.

displacement or by lowering wage rates. For instance, the influx of foreign mathematicians after the Soviet Union collapse temporarily displaced American mathematicians in the academic labor market (Borjas and Doran, 2012). Evidence from Borjas (2005) showed that the inflow of foreign doctorate students had negative impacts on the earnings of native Ph.D graduates within the same field. Further, a series of studies discussed the substitutability of immigrants and natives through examining their wage rate relationship (e.g. John and Zimmermann, 1994; Borjas, 2003; Islam, 2009; Ottaviano and Peri, 2012, etc.).

Motivated by these previous studies, we exploit the impact of exogenous variation in immigration policy on high-skilled labor markets through examining the job placement preference of foreign-born Ph.D graduates. Our analysis is useful for assessing immigration policy effects on high-skilled immigrants and for understanding subsequent complementary/displacing effects that immigrants have on their native counterparts. However, the majority of previous immigration research concentrates on US domestic workers. The effects of immigration policy changes on foreign high-skilled professionals has not receive much attention until recently. Lan (2012, 2013) focused on the effect of Chinese Student Protection Act of 1992 on Chinese Ph.D postdoctoral placement, wage rates, and the number of patents subsequently filed in the United States. Kato and Sparber (2013) investigated the relationship between the variation generated from the reduced H-1B visa cap around 2003 and the SAT scores of international college applicants. Later, Shih (2016) examined the effect of this visa cap change on college enrollment of international students.

The most thematically relevant study to our paper is Amuedo-Dorantes and Furtado (2017). The authors leveraged the differentiation in the availability of H-1B substitute visas for students from different countries and explored how the visa cap reduction affected foreign individuals jobs in academia. The authors use the National Survey of College Graduates (NSCG) data source to focus on the job market for science and engineering(S&E).[4]

However, our research design and data provide a distinct perspective for a broader study population. We focus on how the relaxation in the immigration policy affects foreign-born doctoral

[4]As one of the three surveys in the Scientists and Engineers Statistical Data System (SESTAT surveys), National Survey of College Graduates (NSCG) has its target population focusing on the S&E workers.

recipients' selection into the academic market. From the design point of view, our article studies the H-1B cap exemption regulated from the AC 21 policy. We employ the variation in nationality (domestic/foreign) to examine intent-to-treat effects. More importantly, our study employs licensed data from a *census* of Ph.D holders of all academic majors. Using data of Ph.D recipients from all academic fields is useful to understand foreign professionals' selection into different academic labor markets.

A significant amount of literature explores the H-1B program influence normally using a restricted sample of individuals within STEM fields (e.g. Kerr and Lincoln, 2010; Mithas and Lucas Jr, 2010; Hunt, 2011; Peri, Shih, and Sparber, 2014a,b; Kerr, Kerr, and Lincoln, 2015; Hunt, 2015; Peri, Shih, and Sparber, 2015, etc.). We add to the literature on H-1B program analysis by extending it to all academic fields and achieving better external validity than previous studies.

3.3 Visa Policy and Job Choices

AC 21 passed in January 2000, and it was signed into law in October of the same year by President Clinton. It intended to enhance the effectiveness and portability of the H-1B visa program. The key provision relevant to our study is the special rule allowing universities and research institutions to hire as many foreign-born graduates as they need without being subjected to the annual H-1B quotas.

While there are no specific educational level requirements in AC 21, we argue that the primary treatment effect would be on foreign nationals with doctoral degrees. The favorable visa conditions of AC 21 would potentially motivate all noncitizen graduates to favor academic and research institutions more than other employment prospects. However, academic positions normally require applicants to have a doctoral degree and a strong research background. In this sense, the uptake of the treatment induced by AC 21 would come first and foremost from doctorate recipients. In the following sections, we identify the primary treatment effect on all Ph.D graduates and exploit possible heterogeneous effects by academic fields.

Aside from the annual H-1B quota, foreign-born graduates also face many other constraints when pursuing a job in the United States. After graduation, they are required to file a 12-month

Optional Practical Training (OPT) certificate while applying for the H-1B working visa. Failing to receive an H-1B visa before the OPT expires leaves the prospective worker with a 60-day grace period to leave the country.

In addition, the extra cost incurred during the H-1B application may be significant for small businesses, considering that foreign professionals are paid at the same level as domestic workers. In some cases, in order to hedge the additional costs, companies may offer lower salaries to H-1B recipients compared to American citizens. Moreover, businesses may have less incentives to invest in professional development, training and promotion opportunities for H-1B workers, since they might be mobile if the employer can not provide further sponsorship for a legal permanent residence application. [5]

The enactment of AC 21 significantly reduces potential losses due to a visa rejection, and to a great extent helps foreign-born graduates to avoid the above mentioned constraints, as long as they commit to work in a cap-exempt research institution. Accordingly, the entry barriers become relatively higher for working in industry after the adoption of AC 21. Taking together, jobs in research related institutions may become more attractive after the AC21 policy intervention.

On the other hand, the disadvantages of working in academia, nonprofit and government research institutions are not negligible. The income in these institutions tends to be generous, but it is usually lower than in industry, especially at the entry level. Potential promotions or large salary increases are unusual during the first several years of working in universities and research institutions.

The highly competitive market situation for jobs in academia and research institutions also indicates that the most qualified professionals would probably have opportunities in both academia and industry. The exogenous shock generated from AC 21 will exercise most of its influence on the top qualified foreign doctoral graduates who are potential candidates to both job markets. We hypothesize that with lower barriers to enter the US academic job market, AC 21 may induce individuals to choose a career in H-1B exempted institutions.

[5] http://money.cnn.com/2014/07/30/smallbusiness/immigrant-tech-canada/index.html

Table 3.1: Descriptive Statistics

	Temporary Residents			US Citizen		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Total number of graduates	82136			293386		
Graduates with definite plans	55768			211517		
Demographics						
Male	40382	0.724	0.447	108657	0.514	0.500
Age: 16-25	731	0.013	0.114	1224	0.006	0.076
Age: 26-40	50563	0.916	0.278	145046	0.690	0.463
Father with at least college degree	32429	0.587	0.492	120050	0.572	0.495
Mother with at least college degree	22413	0.406	0.491	97298	0.462	0.499
Single in marital status	19874	0.358	0.480	57218	0.272	0.445
White	14109	0.254	0.435	175772	0.834	0.372
Asian	37042	0.666	0.472	9883	0.047	0.211
Hispanic	2577	0.046	0.210	8922	0.042	0.202
Graduated in public school	38519	0.691	0.462	144805	0.685	0.465
Broadly Classified Doctoral Fields						
Agriculture	567	0.020	0.139	2919	0.020	0.139
Biological/Biomedical Sciences	1011	0.035	0.184	7043	0.047	0.212
Health Sciences	724	0.025	0.157	6809	0.046	0.209
Engineering	10856	0.378	0.485	13260	0.089	0.285
Computer and Information Sciences	2260	0.079	0.269	2834	0.019	0.137
Mathematics	1589	0.055	0.229	2703	0.018	0.133
Physical Sciences	2572	0.090	0.286	8454	0.057	0.231
Psychology	397	0.014	0.117	13115	0.088	0.283
Social Sciences	3312	0.115	0.319	14288	0.096	0.294
Humanities	1889	0.066	0.248	24554	0.165	0.371
Education	870	0.030	0.171	41469	0.278	0.448
Business Management/Administration	2042	0.071	0.257	5373	0.036	0.186
Communication	328	0.011	0.106	2254	0.015	0.122
Fields Not Elsewhere Classified	312	0.011	0.104	4079	0.027	0.163
Outcomes						
Working in Academia	11004	0.383	0.486	80145	0.537	0.499
Working in Industry	15706	0.547	0.498	24378	0.163	0.370

Source: Survey of Earned Doctorates, 1995 to 2006.
"Total graduates" represents Ph.D holders who intend to stay in the US within the next year after graduation.
"Graduates with definite plan" represents individuals with determined study or employment plans and intend to stay in the US for the following year.

3.4 Data

We employ licensed data from the Survey of Earned Doctorates (SED) to execute the quantitative analysis. The SED is a unique annual census conducted since the 1950s. It covers Ph.D graduates from U.S. universities across all fields of study. The survey is implemented at the time of graduation, and hence the response rate is over 90%. We extract micro level data from 1995 to 2006. This time period covers the population of Ph.D graduates who may have been affected by the implementation of AC 21. Data after 2006 is not included in order to construct a clean identification and to rule out potential confounding factors generated from the Great Recession of 2007-2009. The SED also provides detailed information of individual level characteristics, immigration status and post-graduation employment. Since the exogenous shock from AC 21 occurred in 2000, we draw data from six years before and six years after the visa policy change.

The descriptive statistics of the data are shown in Table 3.1. Our analysis is restricted to Ph.D graduates intending to stay in the U.S. after graduation. From the demand side, whether an individual is willing to stay in the United States is affected by the job market conditions of the United States, their home countries and other countries, since holding a doctoral degree from an English speaking country potentially opens the door to global placement. We assume these conditions collectively are exogenous at the time of graduation for a doctorate student. From the supply side, significant world events around the time of the implementation of AC 21 would most likely randomly affect the job preference for academia or industry.

As seen, around 70% of the doctorate recipients, both citizens and foreign-born, have definite plans for employment, study or postdoctoral training after graduation. Of these individuals, half have definite employment plan. Starting from 2004, SED only collect the data of the prospect employers for individuals who indicate they do have a definite work plan. SED does not provide follow-up information regarding the placement of individuals who indicate they do not have definite work plans. Accordingly, we implement the causal analysis with sample restricted to individuals with definite post-graduation plans.[6] If most of the remaining doctorate recipients, who

[6]The Survey of Doctorate Recipients (SDR) collects additional individual level data on post-graduation plans, but

did not have definite employment plans at the time of the survey end up working in academia (industry), then our estimated effect on academic job choice is underestimated (overestimated).

The empirical identification is based on individual level variation in the visa status. Ph.D holders with temporary visas are the most affected cohort and serve as the treatment group, while their US citizen counterparts serve as the comparison group.[7] In the top panel of Table 3.1, we summarize the demographic characteristics of temporary residents and U.S. citizens. The two cohorts differ in many aspects. Males account for 72% of Ph.D recipients with temporary visas while only for 51% of U.S. citizens. By nature, the two groups also show substantial racial differences. The association of interest between the policy change and the job market choice could potentially be driven by treatment and control differences in socioeconomic characteristics. Spurious causal relations can be generated if we fail to adequately control for these observable differences. Thus, we control for individual level demographic information in our extended model specifications.

Individuals from different fields of study face different labor market conditions. Potential heterogeneity across fields of study may also differ across the treatment and comparison groups. The second panel in Table 3.1 shows the differences between foreign and U.S. citizen Ph.Ds in each academic field. The statistics present for the sample focused on individuals with definite employment plans. The majority of native doctorate recipients with earned their doctorates in Humanities and Education; whereas most foreign graduates hold doctorates in Engineering. Effectively controlling the divergence in preferences for doctoral fields is also critical in capturing the potential causal effects from AC 21. To exclude the possibility that the causal effects are driven by changes in the labor market conditions of some specific fields, we allow for flexible interactions of field and time indicators, and explore potential heterogeneous effects by each doctoral field.

Finally, the outcome variables are presented in the last panel of Table 3.1. "Working in

only for a stratified subsample of science and engineering Ph.D.

[7]We exclude the cohort of permanent residents from the analysis for two reasons. First, there is an endogeneity concern of visa status among foreign-born Ph.D recipients. There may be some underlying factors that affect both their job preferences and visa status (e.g. unobserved family backgrounds). It is also possible that the permanent visa status was obtained after earning their Ph.D due to unobserved individual skills and research experience or due to a green card lottery procedure. Both of these factors can drive respondents to favor or reject an academic position. The second concern is the parallel assumption. We test this assumption empirically and find that it is not satisfied for job choices neither in academia nor industry labor markets for green card holders.

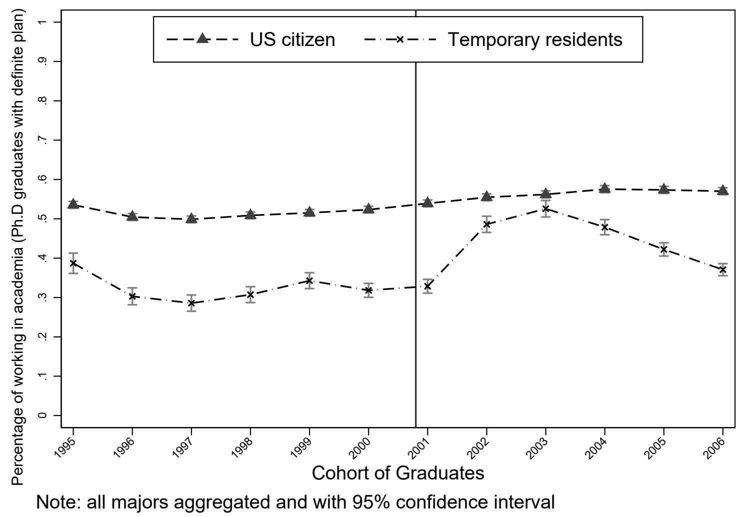
academia" is defined as working in a 4-year college or university, medical school, university-affiliated research institute or community college, which are the options offered in the questionnaire. In addition to working in academia or industry, other respondents stated they would be working in federal/local/foreign government or self-employed. Although the three types of cap exempt entities stipulated in AC 21 involve "government research institutions", there is no further breakdown under this category in the data source.[8]

3.5 Empirical Strategy

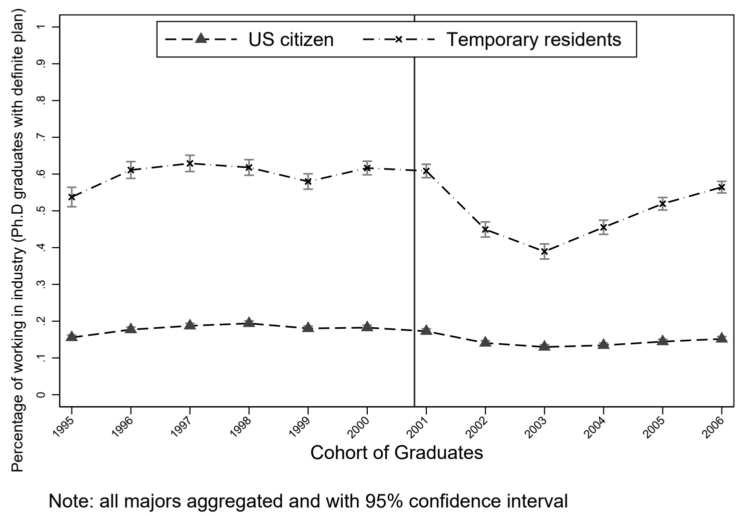
It is possible that the change in the outcome variables between native and foreign Ph.Ds could be driven by inherent differences in the job market conditions of different fields of study. To isolate the causal effect of AC 21 from other potential confounding factors, we implement a Difference-in-Difference estimation using major and academic field fixed effects. That is, we examine whether the probability of working in academia/industry differs more over time for foreign Ph.Ds than domestic Ph.Ds mainly through within-field/major comparisons.

Figure 3.1 plots the time trends of the two outcome variables for all doctoral fields combined. The outcome variables for domestic and foreign-born cohorts follow a similar time trend before the policy intervention. After the policy was implemented (vertical line), there is a sizable shifting pattern for foreign Ph.Ds but not for domestic Ph.Ds. This provides suggestive evidence of an impact from the AC 21 policy intervention. Within five years of the implementation of AC 21, this external shock stimulates a significant increase in the probability of working in academia and a decrease in the proportion of foreign Ph.Ds working in industry.

[8]Working in government as a researcher could be inferred from the database; however, it is not equivalent to working in a government research institution. Note that self-employed is not permitted for foreign nationals. Hence, in order to have a clean identification strategy, we concentrate on the potential causal effects of AC 21 on academia and industry.



(a) Academia



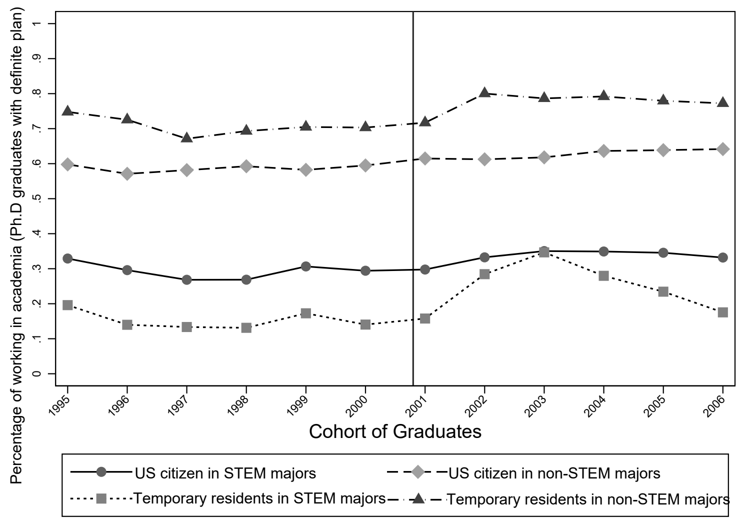
(b) Industry

Figure 3.1: Time Trend of the Share of Graduates Working in Academia

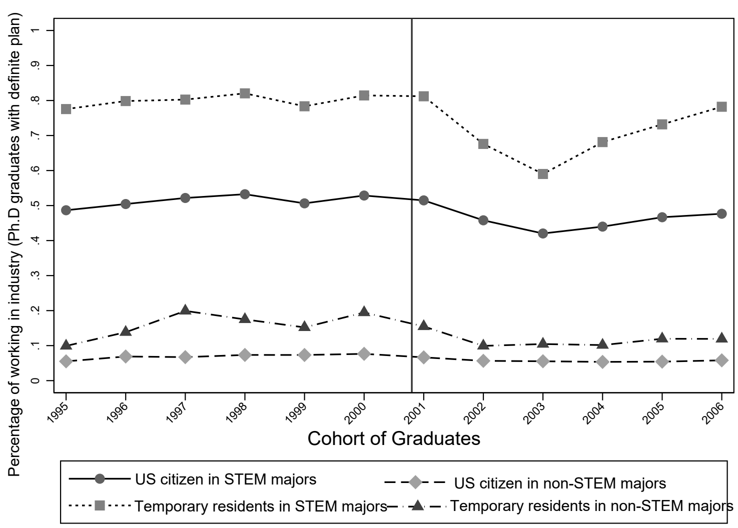
We further disaggregate academic fields into STEM and non-STEM categories and plot the time trends separately in Figure 3.2.[9] Panel (a) shows the placement trajectory of academic jobs,

[9]To leverage the potential variation of job preferences and immigration risk confronted by foreign Ph.Ds in different academic fields, we define STEM majors according to the U.S. immigration and customs enforcement (ICE). Specifically, the STEM categorization we apply is the one used by ICE to regulate which majors are eligible for the revised OPT extension announced in 2008.

and panel (b) for industry jobs. Consistent with our previous findings, both foreign STEM and non-STEM majors follow different time trends after AC 21. The effects for foreign graduates in STEM fields are more pronounced than non-STEM majors. This result further supports our hypothesis of heterogeneous effects across academic fields.



(a) Academia



(b) Industry

Figure 3.2: Time Trend of the Share of Graduates Working in Academia, By STEM Field

The key identifying assumption of the DID strategy is that in the absence of AC 21, the treatment and comparison groups would have experienced similar changes in the probability of working in academia or industry. Visual tests for Figures 3.1 and 3.2 suggest that this assumption is satisfied. To quantitatively capture the treatment effect on job market choices, we establish our estimation model using individual level variation in visa status and graduation time, before and after the policy shock:

$$Y_{ifmt} = \gamma_f + \omega_m + \lambda_t + \boldsymbol{\delta} D_{ft} + X_{ift}\beta + \epsilon_{ifmt} \tag{3.1}$$

where the outcome variable Y_{ift} is a binary indicator equal to 1 if individual i with foreign nationality f obtained his Ph.D degree in year t and worked in academia after graduation, or zero for having an industry job after graduation. The parameters γ_f, ω_m and λ_t are nationality, major and year fixed effects, which capture all the unobserved variations in the outcome variable over citizenship, major and year. X_{ift} are individual characteristics, including race, gender, parent's education, marital status, and university characteristics.[10] D_{ft} is our primary measure for the treatment variable, and it is equal to 1 for foreign Ph.Ds who graduated after the adoption of AC 21. The related parameter $\boldsymbol{\delta}$, indicates the estimated treatment effect. Standard errors are clustered at the fourteen broadly classified academic fields.[11] Due to idiosyncratic features of each doctorate academic field, our baseline specification includes both major and year fixed effects. In order to allow for different disciplines to have distinct trajectories over time, we extend equation 3.1 to include major-specific linear trends. We also estimate specifications that incorporate field-by-year fixed effects to further allow majors of different fields to accommodate differential shocks over time. The results from these specifications and the inclusion of demographic variables serve as augmentations and robustness checks for the baseline results. If the estimated treatment effect changes drastically in any of these extensions, it would indicate that either socioeconomic factors or idiosyncratic trends explain the outcome effects and the baseline parameter may not be associ-

[10]We exclude demographic variables in the baseline model.

[11]We follow the classification in the SED data set to define the broad doctoral academic fields. The list of fields is presented in Table 3.1.

ated with the exogenous shock that we intend to identify. As our point estimates shown in Section 3.6, this is not the case and the additional models support the general results.

3.6 Results

3.6.1 Empirical Results

Our baseline estimates for academic and industry job markets are shown in column (1) of Tables 3.2 and 3.3. Following the baseline results are the estimates for extended specifications which are shown in columns (2) to (4) of these two tables. In addition to the DID linear probability model, we also use a DID probit specification to estimate Equation 3.1. Estimates for the two models are shown in panel A and B of these tables respectively. The results for the linear estimation shown in Table 3.2 indicate that after implementing AC 21, foreign born Ph.Ds are 5% more likely to work in academia than domestic Ph.D graduates. The marginal effect for the probit specification implies an even larger influence, with foreign Ph.D graduates being 7% more likely to work in academic jobs. Our point estimate is robust to the inclusion of controls for idiosyncratic shocks by doctoral fields over time in column (2) and controls for demographic characteristics in column (3). In column (4), we further incorporate major-specific linear time trends into the model, in order to allow for each major to follow a differential trajectory. The parameter of interest remains unchanged. These point estimates are robust and significant at the 1% level throughout all specifications. The results provide suggestive evidence that pre-post variation in job market preferences over cohorts of different visa status is not correlated with other determinants of the outcome variable. Thus, we argue that the estimated effect is driven by the AC 21 policy intervention rather than other factors.

Similarly, Table 3.3 displays analogous estimates for industry job choice. Across all specifications, the estimated parameters are statistically significant at the 1% level. The results show that following AC 21, foreign Ph.D graduates are less likely to work in industry than before the policy intervention. The magnitude of the reduction in probability is around 3 - 4%. The absolute value of the point estimate decreases slightly when we include socioeconomic control variables and other controls for differential trends by doctoral fields and majors.

Table 3.2: Treatment Effect of AC 21 on Job Placement in Academia

Variable	(1)	(2)	(3)	(4)
A. Linear Difference in Difference Estimation				
Noncitizen Ph.Ds Graduated after AC 21	0.047	0.057	0.047	0.047
	(0.002)***	(0.012)***	(0.011)***	(0.011)***
	[0.017]***	[0.019]***	[0.018]***	[0.016]***
Observations	177883	177883	173409	173409
B. Probit Difference in Difference Estimation				
	0.076	0.083	0.073	0.073
	(0.014)***	(0.016)***	(0.015)***	(0.015)***
Observations	177861	177861	173388	173388
Year and Major Fixed Effects	Yes	Yes	Yes	Yes
Field-Year Fixed Effects	No	Yes	Yes	Yes
Major-Specific Linear Time Trends	No	No	No	Yes
Demographic Variables	No	No	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Each cell indicates a separate regression. The dependent variable is whether individual response to place in academia at the time of answering the questionnaire. The explanatory variable of interest refers to noncitizen Ph.Ds who graduated after the adoption of AC 21. Data is restricted to doctorate recipients who have determined plan for employment and intend to stay in the US in the immediate following year. Field-Year fixed effect controls for 385 doctoral majors. Robust standard errors clustered at fourteen broadly classified doctoral fields are shown in parentheses. Following the spirit of Cameron et al. (2011), we additionally implement multi-way clustering at year (12 clusters) and broad doctoral field (14 clusters) for the linear probability model and these standard errors are shown in brackets.

Table 3.3: Treatment Effect of AC 21 on Job Placement in Industry

Variable	(1)	(2)	(3)	(4)
A. Linear Difference in Difference Estimation				
Noncitizen Ph.Ds Graduated after AC 21	-0.054	-0.047	-0.041	-0.038
	(0.010)***	(0.011)***	(0.011)***	(0.011)***
	[0.010]***	[0.015]***	[0.018]**	[0.015]**
Observations	177883	177883	173409	173409
B. Probit Difference in Difference Estimation				
	-0.032	-0.031	-0.028	-0.025
	(0.007)***	(0.007)***	(0.007)***	(0.007)***
Observations	177578	177578	173113	173113
Year and Major Fixed Effects	Yes	Yes	Yes	Yes
Field-Year Fixed Effects	No	Yes	Yes	Yes
Major-Specific Linear Time Trends	No	No	No	Yes
Demographic Variables	No	No	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Each cell indicates a separate regression. The dependent variable is whether individual response to place in industry at the time of answering the questionnaire. The explanatory variable of interest refers to noncitizen Ph.Ds who graduated after the adoption of AC 21. Data is restricted to doctorate recipients who have determined plan for employment and intend to stay in the US in the immediate following year. Field-Year fixed effect controls for 385 doctoral majors. Robust standard errors clustered at fourteen broadly classified doctoral fields are shown in parentheses. Following the spirit of Cameron et al. (2011), we additionally implement multi-way clustering at year (12 clusters) and broad doctoral field (14 clusters) for the linear probability model and these standard errors are shown in brackets.

It suggests that the assessment of AC 21 on the job preference in industry is impacted by variations in doctoral field/major trends and demographic characteristics. This is plausible since the direct effect of AC 21 is primarily exercised on job choice in academia, and hence the inclination of working in industry assumes an indirect effect. Meanwhile, opportunities to work in government and self-employment for native graduates might also have spillover effects on the probability of foreign Ph.Ds to work in industry. These unobserved influencing elements are proxied by socioeconomic and field heterogeneity covariates. Accordingly, additionally controlling for these variables has impact on the magnitude of the point estimates for industry jobs.

3.6.2 Dynamic Difference-in-Difference Estimation

We perform further analysis using the dynamic Difference-in-Difference estimates in the spirit of Cheng and Hoekstra (2013). Specifically, the treatment and control difference of the outcome variable is estimated two/three years prior to the treatment (one year prior to treatment, the year of treatment, one year after treatment, two years after treatment and three years after treatment) relative to the average treatment and control difference four years or more prior to the treatment. By allowing for dynamic divergence between the treatment and control groups over time, this procedure detects how the treatment effect evolves over time. The identification design is invalidated if a significant jump in the dynamic DID coefficients is observed before the year of the treatment. That is, the possibility of other external factors driving the estimated effect cannot be excluded if the highest departing pattern did not appear immediately after the treatment.

Based on Equation 3.1, we build our dynamic Difference-in-Difference estimation model as:

$$Y_{ifmt} = \gamma_f + \omega_m + \lambda_t + \sum_{k=t-3}^{t+3} \boldsymbol{\delta}_k D_{ft} + X_{ift}\beta + \epsilon_{ifmt} \tag{3.2}$$

where the associated coefficients of D_{ft} are the dynamic DID estimates.

The dynamic DID results are shown in Figures 3.3 and 3.4 for academia and industry respectively. We include additional control variables for demographic factors and allow for academic field specific shocks from the baseline model. For both outcome variables of interest, there is no evidence of a significant jump before the enactment of AC 21. The average pre-period coefficient is about 0.6% for academic jobs and 0.5% for industry jobs. Following the passage of AC 21, we observe notable differences in the estimated job preference for both job markets. The magnitude is as large as 9% increase at the peak in academia and 8% decrease at the dip in industry jobs. Throughout the post AC 21 period, the average treatment effect for the academic job market amounts to 5.6% and -4.1% for industry. These results are consistent and almost identical to the traditional DID estimates.

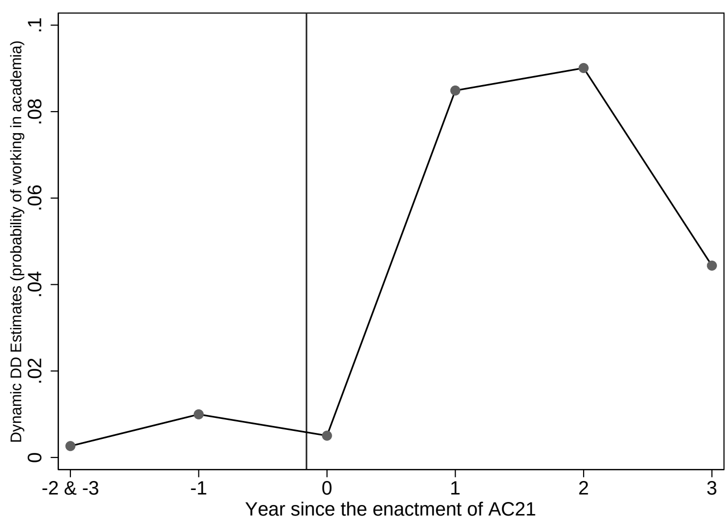


Figure 3.3: Change in the Share of Graduates Working in Academia, Relative to the Difference in Four or More Years Before AC 21

3.7 Robustness

3.7.1 Inference

In Tables 3.2 and 3.3, we show the robust standard errors (in parentheses) clustered at the academic field level. To further investigate the robustness of the results, we implement two additional strategies on inference. The first one is multi-way clustering following Cameron et al. (2011). We specify the robust standard errors clustered at the year and academic field level and re-estimate the

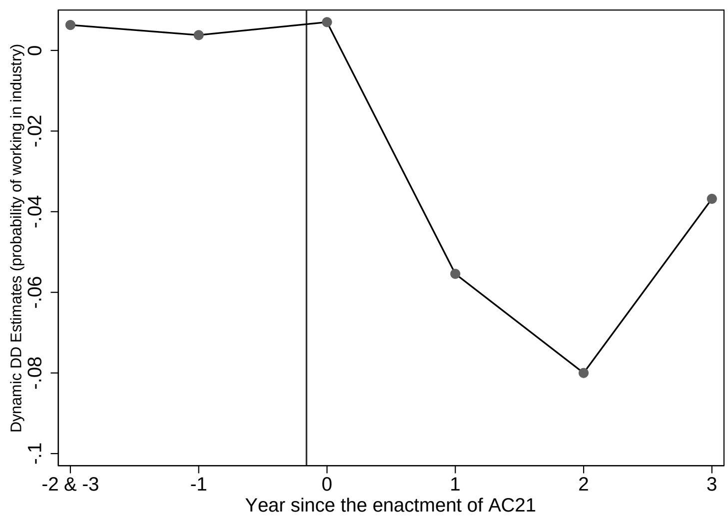


Figure 3.4: Change in the Share of Graduates Working in Industry, Relative to the Difference in Four or More Years Before AC 21

regression across all the specifications. As shown in the brackets of these two tables, the statistical significance does not change at all for the outcome variable of academic job preference. The significance level changes slightly in columns (3) and (4) of Table 3.3 for industry jobs. Consistent with our previous discussion (See section 3.6.1), this again suggests that the job preference in industry assumes a less direct effect from AC 21, and it is sensitive to the inclusion of observed and unobserved factors proxied by demographic characteristics.

The second approach is to construct *p*-values using the bootstrap *t*-procedure (Cameron et al., 2008). With a small number of clusters, the cluster-robust standard errors might be downward biased, this in turn can cause over-rejection of the no-effect null hypothesis. Thusly, we conduct the bootstrap *t*-procedure with 999 replications and present the bootstrapped *p*-value in Table 3.4. As shown, there are no meaningful changes in the statistical significance of the results.

3.7.2 Placebo Experiments

If our major results are affected by external changes other than AC 21, then significant changes may be expected even before the actual adoption of AC 21. To rule out this possibility, we check if the DID estimates are significant to a randomly assigned treatment year before the adoption of

Table 3.4: Inference for Treatment Effect of AC 21: Wild Bootstrap-t Procedure

Linear Difference in Difference Estimation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Academia	Academia	Academia	Academia	Industry	Industry	Industry	Industry
Parameter estimates	0.047***	0.057***	0.047***	0.047***	-0.054***	-0.047***	-0.041***	-0.038***
P-value								
Cluster-robust	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
Wild cluster bootstrap-t	0.002	0.000	0.004	0.000	0.000	0.002	0.006	0.008
Year and Major Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field-Year Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Major-Specific Linear Time Trends	No	No	No	Yes	No	No	No	Yes
Demographic Variables	No	No	Yes	Yes	No	No	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Each cell indicates a separate regression. The dependent variable is whether individual response to place in industry (for the first four columns) and academia (for the second four columns) at the time of answering the questionnaire. The explanatory variable of interest refers to noncitizen Ph.Ds who graduated after the adoption of AC 21. Data is restricted to doctorate recipients who have determined plan for employment and intend to stay in the US in the immediate following year. Field-Year fixed effect controls for 385 doctoral majors. Standard errors are clustered on fourteen broadly classified doctorate fields and we perform wild cluster bootstrap-t test with 999 replications, following an example in Cameron et al. (2008).

AC 21. Specifically, we conduct several placebo exercises to compare the outcome differences in the treatment and comparison groups over different designs of pre- and post-treatment year arrangements.

Columns (1) to (4) of Table 3.5 show the results of placebo tests over different specifications for academic job preference. Columns (5) to (8) display the equivalent estimation for industry job choice. Each row represents a different pre/post arrangement scenario. For example, the first row assigns the placebo intervention year to 1995. Across the estimation scenarios and specifications, nearly all of the estimates indicate no effect. It is noteworthy to detect no significant effects in the placebo tests given the large sample size of the dataset used in this study.

3.7.3 Falsification Test

The identification of the Difference-in-Difference strategy is based on two key assumptions. The first assumption is the parallel trend assumption for foreign-born and U.S. domestic graduates. Namely, without the policy intervention, the treatment and comparison group would have undertaken the same time trend over time. We checked the first assumption qualitatively and quantitatively in previous sections and presented the results in Figures 3.1 to 3.4. The other assumption states that except for the exogenous shock generated from the policy intervention, other external variations are assumed to be orthogonal to the outcome of interest after effectively controlling for observable differences.

In this section, we will test this second assumption. Constructing a test for the second assumption in our study is equivalent to checking whether Ph.D graduates who are exogenous to this policy intervention also change their job type preference after its implementation.

Foreign-born Ph.D graduates who are trained as post-doctoral scholars are not required to change their visa status from a F1 student visa to a H-1B work visa. Throughout the duration of the Optional Practical Training (OPT), foreign Ph.D holders can acquire post-doctoral training under the same student visa status. Therefore, individuals seeking a post-doctoral position are not affected by the AC 21 policy change. There are no additional incentives inducing would-be-postdocs to change their plan towards employment in academia. Moreover, foreign Ph.Ds who are

Table 3.5: Placebo Effect on Job Placement in Academia and Industry in Pre-periods, Linear DID Estimation

Placebo intervention happened at:	(1) Academia	(2) Academia	(3) Academia	(4) Academia	(5) Industry	(6) Industry	(7) Industry	(8) Industry
1995	-0.023	-0.016	-0.019	-0.018	0.026	0.021	0.022	0.022
	(0.022)	(0.020)	(0.020)	(0.020)	(0.022)	(0.021)	(0.023)	(0.021)
1996	-0.004	-0.001	-0.003	-0.003	0.015	0.016	0.018	0.019
	(0.015)	(0.013)	(0.012)	(0.012)	(0.020)	(0.017)	(0.017)	(0.015)
1997	0.008	0.009	0.006	0.006	-0.000	0.003	0.006	0.006
	(0.011)	(0.012)	(0.011)	(0.012)	(0.014)	(0.012)	(0.010)	(0.010)
1998	0.019*	0.017*	0.012*	0.012	-0.013	-0.011	-0.004	-0.005
	(0.009)	(0.009)	(0.007)	(0.007)	(0.014)	(0.011)	(0.009)	(0.007)
1999	0.008	0.011	0.009	0.009	-0.002	-0.001	0.001	0.002
	(0.009)	(0.010)	(0.009)	(0.010)	(0.017)	(0.013)	(0.012)	(0.009)
Observations	89188	89188	87187	87187	89188	89188	87187	87187
Year and Major Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Field-Year Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Major-Specific Linear Time Trends	No	No	No	Yes	No	No	No	Yes
Demographic Variables	No	No	Yes	Yes	No	No	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Each cell indicates a separate regression. Suppose treatment happened in 1995, 1996, 1997, 1998 or 1999, we test whether there is significant jump even before the actual AC 21 (enacted at the end of 2000) took into effect. For instance, we set pre-period from 1995 to 1997, post-period from 1998 to 2000 for placebo intervention supposed to happen at the end of 1997.

The dependent variable is whether individual response to place in academia at the time of answering the questionnaire. The explanatory variable of interest refers to noncitizen Ph.Ds who graduated after the adoption of AC 21. Data is restricted to doctorate recipients who have determined plan for employment and intend to stay in the US in the immediate following year. Field-Year fixed effect controls for 385 doctoral majors. Robust standard errors are clustered at fourteen broadly classified doctoral fields.

qualified for a high standard faculty position would normally not be influenced by this visa policy to take a post-doctoral position instead of a full-time faculty opportunity. It is also reasonable to assume that a foreign doctorate holder with a permanent job offer in industry would not be driven by AC 21 to favor a post-doctoral position. Essentially, AC 21 only affects the job preference of individuals with job prospects subject to the regulation of the H-1B policy. Due to this feature, AC 21 would not have a significant impact on the aggregate participation level of post-doctoral scholars. In order to exclude other external events taking place around the same time of AC 21 driving the results in Section 3.6, analogous estimations for post-doctoral scholars, unaffected by the AC 21, are implemented in this section and they should yield null effects.

According to this line of reasoning, we perform similar estimations to post-doctoral participation. Table 3.6 presents the point estimates for our main specifications using DID linear probability and DID probit models. Across columns (1) to (4), we gradually augment our specifications to allow for idiosyncratic shocks by academic fields over time, each doctoral major to follow their own time trend and controlling for individual level characteristics. Overall, the sign of the parameters are unstable and there are no statistically significant effects of AC 21 on post-doctoral participation. As the data source used for the analysis is the census of the entire population, this falsification tests provides convincing support to our main findings and rule out other factors driving the effects identified in Section 3.6.

It is possible that the post-doctoral job market participation may be stable enough to remain unaffected by most exogenous variations. If this is the case, the detected orthogonality of post-doctoral participation to AC 21 might not provide sufficient evidence about the exclusion of other possible influencing shocks. However, a report from the National Science Foundation (NSF), estimated that the number of post-doctoral researchers had almost doubled from 1987 to 2002 (Thurgood, 2004). Lan (2012) documents a 24% decrease in post-doctoral participation among Chinese graduates in response to an unexpected increase of the permanent visa granting due to the Chinese Student Protection Act of 1992. Stephan and Ma (2005) employed the same data set used in our study to examine the underlying factors for the rise of post-doctoral participation. All these

Table 3.6: Falsification Tests: The Effect of AC 21 on Postdoctoral Participation

Variable	(1)	(2)	(3)	(4)
A. Linear Difference in Difference Estimation				
Noncitizen Ph.Ds Graduated after AC 21	0.009	-0.002	-0.003	0.002
	(0.013)	(0.008)	(0.008)	(0.007)
Observations	264636	264636	258042	258042
B. Probit Difference in Difference Estimation				
	-0.002	-0.006	-0.007	-0.004
	(0.009)	(0.007)	(0.007)	(0.006)
Observations	264518	264518	257771	257771
Year and Major Fixed Effects	Yes	Yes	Yes	Yes
Field-Year Fixed Effects	No	Yes	Yes	Yes
Major-Specific Linear Time Trends	No	No	No	Yes
Demographic Variables	No	No	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Each cell indicates a separate regression. The dependent variable is whether individual response to postdoc participation after graduation. The explanatory variable of interest refers to noncitizen Ph.Ds who graduated after the adoption of AC 21. Data is restricted to doctorate recipients who have definite plan for study or employment and intend to stay in the US in the immediate following year. Field-Year fixed effect controls for 385 regular doctoral fields. Robust standard errors are clustered at fourteen broadly classified doctoral fields.

studies provide support to the notion that the unresponsiveness of post-doctoral participation to AC 21 is driven by its orthogonality to the policy change.

3.8 Heterogeneous Effects by Doctoral Field

In this section we investigate whether AC 21 affects foreign Ph.Ds asymmetrically by academic field. As shown in Table 3.1 domestic and foreign-born doctoral recipients diverge in their preferences for academic fields. For instance, foreign Ph.D students are more likely to select STEM related majors compared to domestic students. In order to account for this potential selection, we control for detailed major fixed effects and allow for different majors to follow their own linear trend in our aggregate estimation presented in previous sections. Since the supply and demand of the job market conditions for different fields naturally differs, an individual field examination will provide an indication of where the exact treatment effect comes from. Therefore, an investigation on the potential heterogeneous responses of job choices to AC 21 by academic field is needed.

Meanwhile, we argue that the exogenous shock from AC 21 is independent of the field selection of foreign-born Ph.Ds. Considering the strong requirements and specialized information required by doctoral programs, it is unlikely that foreign Ph.Ds would strategically change their fields in response to AC 21. Furthermore, our post-period includes six years after AC 21, covering the full length of most doctoral programs. Most Ph.D cohorts who registered after the full implementation of this policy intervention would also officially graduate after 2006, and these cohorts are already excluded from our data.

We stratify the sample to fourteen broadly classified doctoral fields as defined in Table 3.1, except for the last category "Fields Not Classified Elsewhere". The point estimates for all the other thirteen categories are presented in Tables 3.7 and 3.8. Equation 3.1 is implemented in our DID specification with linear probability and probit models. The heterogeneity in the estimates suggests that our aggregate estimate of career preference in academia is largely driven by Ph.D graduates majoring in Agriculture and Life Science (AGLS), Mathematics and Business Management (BM). The point estimates for these fields are about twice as large as the estimates for Biological Science and Computer Science (CS), and more than three times than that for Engineering and Social Science. Specifically, the point estimate for Computer Science is 59% of that for AGLS in the linear probability model and 53% of that for AGLS in the probit model. The estimated treatment effect for Engineering is even lower, assuming 21% of the point estimate for AGLS in the linear model and 28% of that for AGLS in the probit model . The responses of other fields such as Health Sciences, Physical Sciences and Psychology to AC 21 are positive for choosing academia, but the effects are not statistically significant.

For industry jobs shown in Table 3.8, the heterogeneity in career preferences are mainly driven by AGLS, CS and BM. Overall, we observe more variation in the significance of point estimates for career choices in industry. For instance, Mathematics is only marginally affected by this policy in both models and the parameter significance level for CS decreases to 5% in the linear model. This is consistent with our previous aggregate results in Section 3.6.1 and might be plausibly explained

Table 3.7: Heterogeneous Treatment Effects of AC 21 on Job Placement in Academia, By Doctorate Field

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
A. Linear Difference in Difference Estimation													
Noncitizen Ph.Ds Graduated after AC 21	0.145***	0.052**	0.025	0.031***	0.086***	0.120***	-0.014	0.046	0.036**	-0.022	-0.013	0.091***	-0.010
	(0.037)	(0.022)	(0.019)	(0.011)	(0.006)	(0.030)	(0.014)	(0.076)	(0.013)	(0.016)	(0.033)	(0.026)	(0.026)
Observations	3417	7903	7357	23462	4943	4188	10790	13188	17161	25732	41297	7196	2508
B. Probit Difference in Difference Estimation													
	0.196***	0.063**	0.032*	0.054***	0.104***	0.118***	0.001	0.048	0.035**	-0.025	-0.007	0.085***	-0.003
	(0.044)	(0.026)	(0.018)	(0.009)	(0.003)	(0.029)	(0.021)	(0.083)	(0.015)	(0.025)	(0.044)	(0.021)	(0.027)
Observations	3417	7902	7357	23460	4943	4188	10790	13188	17161	25714	41297	7196	2508
Year and Major Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Major-Specific Linear Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Each cell indicates a separate regression. The dependent variable is whether individual response to place in academia at the time of answering the questionnaire. The explanatory variable of interest refers to noncitizen Ph.Ds who graduated after the adoption of AC 21. Data is restricted to doctorate recipients who have determined plan for employment and intend to stay in the US in the immediate following year. Robust standard errors are clustered at doctorate majors.

Doctoral field corresponding to each column is as the following:

(1) Agriculture and Life Science; (2) Biological/Biomedical Sciences; (3) Health Sciences; (4) Engineering; (5) Computer and Information Sciences; (6) Mathematics; (7) Physical Sciences; (8) Psychology; (9) Social Sciences; (10) Humanities; (11) Education; (12) Business Management; (13) Communication.

Table 3.8: Heterogeneous Treatment Effects of AC 21 on Job Placement in Industry, By Doctorate Field

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
A. Linear Difference in Difference Estimation													
Noncitizen Ph.Ds Graduated after AC 21	-0.068**	0.012	-0.005	-0.005	-0.096**	-0.072*	0.025*	-0.097*	-0.017	-0.006	-0.003	-0.083***	0.010
	(0.027)	(0.027)	(0.033)	(0.012)	(0.014)	(0.036)	(0.013)	(0.053)	(0.011)	(0.009)	(0.017)	(0.026)	(0.023)
Observations	3417	7903	7357	23462	4943	4188	10790	13188	17161	25732	41297	7196	2508
B. Probit Difference in Difference Estimation													
	-0.072***	0.011	0.003	-0.041***	-0.108***	-0.053*	0.021	-0.065*	-0.003	-0.013*	0.001	-0.054***	0.010
	(0.018)	(0.027)	(0.021)	(0.013)	(0.010)	(0.028)	(0.017)	(0.028)	(0.010)	(0.006)	(0.009)	(0.014)	(0.022)
Observations	3417	7902	7357	23460	4943	4188	10790	13170	17161	25431	41297	7196	2499
Year and Major Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Major-Specific Linear Time Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Each cell indicates a separate regression. The dependent variable is whether individual response to place in industry at the time of answering the questionnaire. The explanatory variable of interest refers to noncitizen Ph.Ds who graduated after the adoption of AC 21. Data is restricted to doctorate recipients who have determined plan for employment and intend to stay in the US in the immediate following year. Robust standard errors are clustered at doctorate majors.

Doctoral field corresponding to each column is as the following:

(1) Agriculture and Life Science; (2) Biological/Biomedical Sciences; (3) Health Sciences; (4) Engineering; (5) Computer and Information Sciences; (6) Mathematics; (7) Physical Sciences; (8) Psychology; (9) Social Sciences; (10) Humanities; (11) Education; (12) Business Management; (13) Communication.

by the fact that the provisions of AC 21 only impose an indirect effect on industry job choices.

3.9 Concluding Remarks

The prevalent debate surrounding immigration policy has attracted much attention. One key issue often discussed is whether high-skilled immigrants have a complementary or displacing effect on domestic workers. A crucial challenge for detecting this effect relies on overcoming the endogeneity of the behavior of immigrants in the job market. An evaluation of how immigration policy change influences job market preferences of high-skilled immigrants paves the way to accurately identify the impact of foreign workers on domestic workers in the future. In this study, we leverage the H-1B policy shock generated by the AC 21 as an exogenous variation and capture changes in job choices of foreign Ph.D graduates in academia and industry. AC 21 eliminates the cap on the number of H-1B visas to be granted for applicants in universities, non-profit and government research institutions.

Our findings indicate that by reducing the potential risk in the process of the H-1B petition, AC 21 on average causes foreign Ph.D graduates to be 5% more likely to start a career in academia and indirectly causes them to be 3-4% less likely to begin a career in industry. Our point estimates for the direct treatment effect are robust to the inclusion of various controls including individual level characteristics, differing trends in majors and idiosyncratic shocks on academic fields over time, while the indirect effect is subject to slight sensitivity to the specifications. A falsification test on post-doctoral participation and placebo experiments based on pre-period data further support the estimated results, excluding other possible external changes in the labor market. Heterogeneous examinations indicate that the academic major treatment effect is driven by Ph.D recipients in Agriculture and Life Science, Mathematics and Business Management.

We present evidence that eliminating the H-1B visa cap for research institutions makes newly foreign Ph.Ds more likely to stay in academia in the United States. Meanwhile, these estimated results do not take into consideration the potential growth among foreign Ph.Ds to work in non-profit/government research institutions. To the extent that Ph.D recipients are driven to seek employment in these institutions, our estimated effects will understate the potential impact on academic labor markets.

One notable pattern for our results is that the effects from AC 21 phased out eventually as the Great Recession approached. From the perspective of demand side, as all cap-exempt institutions are non-profit entities, it is plausible that these institutions were not able to fully digest the increased supply of noncitizen Ph.D generated from this policy. Besides the AC 21, there is also a lack of follow-up favoring policies facilitating qualified institutions to boost the demand for noncitizen doctoral recipients. This unbalance was further intensified as the Great Recession finally commenced and we see a trend of career preference back to its previous levels after five years of the adoption of AC 21.

Our findings have critical policy implications. We argue that the accumulated effect from this exogenous variation could be even larger. As identified by previous studies, the influx of high-skilled immigrants, particularly Ph.D holders, makes great contributions to the number of patents filed and the development of fundamental research across academic fields. Admittedly, in the field of scientific research, sometimes contributions from several key figures could bring ground-breaking reform to a discipline. Therefore, the long-term effect from this temporary increase in the supply of doctorate recipients, of all academic fields of study, to the research institutions is also worth examining. Unfortunately, due to data limitations, we were not able to exploit it in this study. In order to understand the full impact of AC 21, future studies can also investigate its effects on wages in high-skilled labor markets.

4. IDENTIFYING MISCLASSIFICATION ERRORS IN SUBJECTIVE WELL-BEING: A NEW APPROACH TO MAP HAPPINESS

> "Happiness is the meaning and purpose of life, the whole aim and end of human existence."
>
> *Aristotle, Nichomachean Ethics, book 1, section 7.*

4.1 Introduction

The pursuit of happiness has been paramount for humankind throughout history. It has evoked the attention of philosophers from Epicurus and Aristotle to Jeremy Bentham and John Stuart Mill. The U.S. constitution declares that the pursuit of happiness is an unalienable right of every individual. But what is happiness? And how can it be measured? A recent and growing literature argues that income and consumption are not ideal measures of well-being (see Fleurbaey (2009) for a review). Subjective well-being (SWB) can arguably be used as a standard for measuring social welfare and progress. The contemporary literature related to the economics of happiness started with the pioneering work of Richard Easterlin (1974). Recently, SWB is gaining attention in economic research and welfare evaluation.

The most frequently used method to elicit SWB is to directly ask people about their feelings of happiness and overall life satisfaction (e.g., Frey and Stutzer, 2000, 2002, 2010; Diener, 2000). This method is attractive due to its simplicity and versatility. However, compared to objective indicators of social welfare, SWB is subject to substantial misreporting or measurement errors. Self-reported measures of happiness can be largely disturbed by contextual factors, and personal reasons such as selective memory and projection bias. Furthermore, people may avoid reporting extreme categories of SWB, deceiving themselves to believe that they are more (or less) happy than they actually are (e.g., Hagedorn, 1996; Sheridan et al., 2015). Individuals may also be easily influenced by random shocks of recent events at the time of responding to surveys due to projection bias (Chetty, 2015). For example, dealing with health problems or the loss of a close relative may negatively impact happiness measures. On the contrary, a recent job promotion or the birth of a child may positively influence self-reported happiness. More importantly, it is difficult for some people to precisely evaluate their subjective well-being. Using a hypothetical survey of SWB rankings under different scenarios, Benjamin et al. (2012) find systematic reversals in individual choices of reported happiness. Although there is evidence of measurement errors in self-reported happiness (OECD, 2013; Diener et al., 2013; Chetty, 2015), this issue has been mostly ignored by empirical practitioners. Since precise and reliable measurements of SWB are fundamental for understanding the economics of happiness, we believe that it is critical to explore ways to recover the true latent status of SWB.

To address this concern, we apply a novel closed-form identification and estimation method first proposed by Hu (2008) and developed in Feng and Hu (2013). In general, the methodology employs an eigenvalue-eigenvector decomposition technique to establish a closed-form identification for the misclassification matrix of SWB, which is constructed a priori using the conditional distribution of reported happiness. By pre-multiplying the inverse of the identified misclassification matrix by the unconditional distribution of SWB, the underlying true distribution of latent SWB is directly estimated. This method is becoming popular in recent empirical applications (e.g., An et al., 2010, 2017). The identification strategy in our article is implemented in the spirit of Feng and Hu (2013).

The data source that we employ to recover the true underlying distribution of latent SWB is the Integrated European/World Value Survey (EVS/WVS). EVS and WVS have been widely used in economic research (e.g., Stevenson and Wolfers, 2008b; Campante and Yanagizawa-Drott, 2015; Stevenson and Wolfers, 2013; Alesina and Giuliano, 2010; Galor and Özaka, 2016). Besides its broad use, we use these data because it contains three direct measures of SWB across countries and time, which are required for identification.[1] For sake of computation, we select 80 countries based on geographical diversity and consistency of survey data across periods.

[1]Other data sources for SWB include: World Happiness Report, Pew Global Attitudes Survey, Gallup World Poll and the International Social Survey Programme.

We use reported "feeling of happiness" in EVS/WVS as the direct measurement of SWB, while "life satisfaction" (LS) and "freedom of choice and control in life" (FCC) serve as two repeated measures of SWB.[2] Under relatively weak assumptions, we impose a structure on the misclassification of reported happiness and are able to obtain the latent SWB probability distributions by country and demographic groups. Based on these results, we exploit country characteristics and specific demographic groups to explain misclassification errors. Particularly, we examine how the prevalent religion of a country affects the magnitude and direction of the bias in self-reported happiness. The relationship between religion and SWB has been abundantly discussed (e.g., Ellison, 1994; Dolan et al., 2008; Deaton and Stone, 2013; Argyle, 2003; Francis, 2010; Campante and Yanagizawa-Drott, 2015). In this article, we add to the literature by examining this association through the lens of identifying misclassification errors in SWB.

Another motivation of our study is related to the well-known "Easterlin paradox" (e.g., Easterlin, 1974, 1995; Blanchflower and Oswald, 2004); See Kahneman and Krueger (2006); Frey and Stutzer (2002); Clark et al. (2008) for a more comprehensive review. *The Easterlin paradox* posits that although there is a positive correlation between individual income and measures of SWB, aggregate happiness is not significantly associated with GDP *per capita*. Recent studies have challenged the validity of the Easterlin paradox and argue that there is robust evidence of a positive relationship between SWB and income across countries and over time (e.g., Hagerty and Veenhoven, 2003; Deaton, 2008; Stevenson and Wolfers, 2008a; Sacks et al., 2010; Stevenson and Wolfers, 2013).

Meanwhile, a modified version of the Easterlin's hypothesis has also been proposed. It argues that the correlation between income and SWB only exists for individuals with income below a certain threshold, beyond which income is no longer related to happiness (e.g., Frey and Stutzer, 2002; Clark et al., 2008; Di Tella and MacCulloch, 2008; Di Tella et al., 2010). For instance, Clark et al. (2008) state that "greater economic prosperity at some point ceases to buy more happiness." The debate about the Easterlin paradox and its modified hypothesis is likely to continue unless

[2]A detailed explanation for the choice of variables is presented in Section 2.3.

SWB can be more precisely measured and quantified.

Our study makes the first attempt to use a sophisticated method to correct misclassification errors in self-reported SWB. Our analysis produced the following major findings. First, self-reported happiness elicited from surveys has substantial misclassification errors. Second, misclassification errors of SWB are heterogeneous across countries with different prevalent religious beliefs. Third, our analysis shows that religious beliefs and economic development impact how misclassification errors relate to demographic characteristics. Lastly, we use the corrected SWB obtained by our latent variable approach to reexamine the Easterlin paradox and modified-Easterlin hypothesis. Although *reported* measures of SWB provide support for the Easterlin hypothesis, using our *corrected* measures of SWB lead to reject the *original* and *modified* Easterlin hypotheses.

The rest of the paper proceeds as follow. Section 2 documents the data. Section 3 introduces the identification strategy. Section 4 reports the results, and section 5 concludes. A complete proof for the theorem presented in section 3 is reported in the Appendix. Summary statistics of the data and other detailed subsidiary results are presented in the supplementary materials.

4.2 Data

The main data source for this study is the Integrated European/World Value Survey (EVS/ WVS hereafter) from 1981 to 2014. There are four waves in EVS and six waves in WVS, providing a wide coverage of countries around the world.[3] The two surveys do not define the years and waves in exactly the same way. For a particular country, the data appears in either the EVS or WVS survey, and for most countries, data were not continuously collected across waves. As a result, the country level data are unbalanced across years. Previous studies, such as Stevenson and Wolfers (2013), employ cross sectional analysis by wave. The potential concern with this method is that each estimation includes different countries. We attempt to make cross country comparisons feasible by pooling available data from all waves by country and identifying the true latent SWB for each

[3]The four waves of EVS data are from 1981-1984, 1990-1993, 1999-2001 as well as 2008-2010, and the six waves of WVS data are from 1981-1984, 1989-1993, 1994-1998, 1999-2004, 2005-2009, 2010-2014. A list of countries and waves is provided in the supplementary material.

demographic group countrywide.[4]

The three key variables selected as the measures for subjective well-being come from 1) level of happiness, 2) overall life satisfaction (LS) and 3) freedom of choice and control in life (FCC). The "happiness" variable is the directly reported measure of SWB (Helliwell et al., 2012; OECD, 2013). We use LS as another close measurement of SWB, as suggested by previous literature (OECD, 2013). FCC is chosen as the third measurement of SWB for identification purposes, which will be further elaborated in the next section.

Since the three SWB questions use different scales, to identify the latent status of SWB, we recoded them to guarantee all three variables of interest share a common support. The variables are defined as follows: 1) Happiness has three levels: "unhappy", "quite happy" and "very happy"; 2) overall life satisfaction can be: "most dissatisfied","quite satisfied" and "most satisfied"; 3) freedom of choice and control can be: "not at all", "quite a bit" and " a great deal". The overall distributions of the three variables are similar, with around 14-19% of the observations in the first category, 51-57% in the second category, and 23-27% in the third category.

In addition to demographic characteristics (e.g., age, gender, family income) provided in the EVS/WVS dataset, country-level characteristics are also used in the analysis. GDP *per capita* was obtained from the Penn World Table (Feenstra et al., 2015). The World Factbook was used to identify the prevalent religion of each country.

4.3 Nonparametric Identification

In order to correct potential under/over-reporting in each category of self-reported subjective well-being, we apply a recently proposed closed-form identification and estimation method in the spirit of Feng and Hu (2013) and Hu (2008) to recover an individual's latent true classification of SWB. Note that Hu (2008) sets up a 2-measurement model of the latent variable, while Feng and Hu (2013) develop a 3-measurement model of the latent variable. Given our data structure, we

[4]Notice that our strategy of identifying only one set of latent SWB status for each country-demographic group combination leveraging all available data over time makes it possible to do cross-country comparisons; however, this strategy also makes our results not compatible with other studies using cross-sectional data wave by wave or panel data from other sources.

employ a 3-measurement model to identify the underlying SWB.

4.3.1 Basic Set-up and Assumptions

Let X be the reported status of SWB, and X^* be the underlying true status of SWB. We use the reported status of *happiness* in the EVS/WVS as the direct measurement of SWB, X. Y and Z represent two other indirect measurements of SWB. Y and Z were obtained from responses to questions related to *life satisfaction* (LS) and *freedom of choice and control in life* (FCC). We further control for several sociodemographic factors including age, gender and income, which in a simple OLS regression contribute significantly to direct measurements of SWB. Identification subsamples are defined based on individual characteristics, represented in the model as $\mathbf{C}$. We observe an i.i.d. set $\{X, Y, Z, \mathbf{C}\}_i$, where $i = 1, ..., N$. The three categories for each SWB measurement variable are listed below:

$$X = \left\{ \begin{array}{cc} 1 & \text{Unhappy} \\ 2 & \text{Quite Happy} \\ 3 & \text{Very Happy} \end{array} \right\}, Y = \left\{ \begin{array}{cc} 1 & \text{Most Dissatisfied} \\ 2 & \text{Quite Satisfied} \\ 3 & \text{Most Satisfied} \end{array} \right\}, Z = \left\{ \begin{array}{cc} 1 & \text{Not At All} \\ 2 & \text{Quite A Bit} \\ 3 & \text{A Great Deal} \end{array} \right\}$$

Next, we define the misclassification probabilities and distributions for the latent SWB status in matrix form. For each demographic group $\mathbf{c}$, one realization of the combination of the three individual characteristics, the misclassification matrix is defined as:

$$\mathbf{H}_{X|X^*,\mathbf{c}} \equiv \left[Pr(X = i|X^* = k, \mathbf{c})\right]_{i,k} \tag{4.1}$$

The diagonal matrix is defined as:

$$\mathbf{H}_{Y=3|X^*,\mathbf{c}} \equiv \quad \left[Pr(Y = 3|X^* = k, \mathbf{c})\right]_{k} \tag{4.2}$$

Similarly, other distribution matrices are defined as:

$$\mathbf{H}_{Y=3,X,Z|\mathbf{c}} \equiv \quad [Pr(Y=3,X,Z|\mathbf{c})]_{i,k} \tag{4.3}$$

$$\mathbf{H}_{X,Z|\mathbf{c}} \equiv \quad [Pr(X,Z|\mathbf{c})]_{i,k} \tag{4.4}$$

$$\mathbf{H}_{X^*,Z|\mathbf{c}} \equiv \quad [Pr(X^*,Z|\mathbf{c})]_{i,k} \tag{4.5}$$

The assumptions required for identification of the underlying true SWB status are provided below.

Assumption 1. $Pr(X|X^*, Z, \mathbf{c}) = Pr(X|X^*, \mathbf{c})$

This assumption implies that conditional on the true SWB status and individual characteristics, the reported status does not rely on other measurements of SWB. The Z measure is only allowed to correlate with misclassification errors through the latent true SWB status. The difference between the latent X^* and X is assumed to be independent with respect to other measures represented by Z.

Assumption 2. $Pr(Y|X^*, X, Z, \mathbf{c}) = Pr(Y|X^*, \mathbf{c})$

Similarly to Assumption 1, Assumption 2 indicates that reported SWB (X) and the measure of Z affect the probability distribution of reported LS only through the latent true SWB status; X or Z does not contain additional useful information on reported LS beyond the true SWB and individual characteristics $\mathbf{c}$. Note that Feng and Hu (2013), by the first-order Markov restriction, essentially assume that $Pr(Y^*|X^*, Z^*, \mathbf{c}) = Pr(Y^*|X^*, \mathbf{c})$. In our case we impose Assumption 2 following Hu (2008).

Assumption 3. *Full rank condition:* $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) = 3$.

We quantitatively test the rank condition using the method proposed by Robin and Smith (2000). The full rank of $\mathbf{H}_{X,Z|\mathbf{c}}$ is tested in a sequential procedure. Specifically, we first test the null hypothesis of rank 1 against rank 2, and then rank 2 against rank 3. The results of the rank test are presented in the supplementary materials. In our setting, to leverage the potential variation between FCC and SWB, FCC was chosen as Z. [5].

Assumption 4. $Pr(Y = 3|X^* = i, \mathbf{c}) \neq Pr(Y = 3|X^* = j, \mathbf{c})$ *for all* $i \neq j$.

Assumption 5. *Given the measure of* Y *and individual characteristics* $\mathbf{c}$*, the conditional probability* $Pr(Y = 3|X^*, c)$ *is strictly increasing in* X^*.

Assumption 3 guarantees the invertibility of $\mathbf{H}_{X,Z|\mathbf{c}}$, and Assumption 4 indicates the three elements in the diagonal matrix $\mathbf{H}_{Y=3|X^*,\mathbf{c}}$ are different. Assumption 5 is equivalent to assuming that when LS status is observed in level 3 (most satisfied with life), it is more likely that individuals make their judgement based on their latent SWB of level 3 as well. Similarly, it is less likely that a self-reported level of 3 in LS is based on a lower underlying SWB (i.e., level 1 or 2).

4.3.2 Estimation Procedures

The identification procedure for latent SWB is constructive. The misclassification matrix will be directly estimated by the following steps. First, by the law of total probability and stated assumptions, the following relationships are satisfied:

$$Pr(Y, X, Z|\mathbf{c}) = \sum_{X^*} Pr(Y|X^*, \mathbf{c}) Pr(X|X^*, \mathbf{c}) Pr(X^*, Z|\mathbf{c}) \tag{4.6}$$

$$Pr(X, Z|\mathbf{c}) = \sum_{X^*} Pr(X|X^*, \mathbf{c}) Pr(X^*, Z|\mathbf{c}) \tag{4.7}$$

In matrix form, we have

$$\mathbf{H}_{Y=3,X,Z|\mathbf{c}} = \mathbf{H}_{X|X^*,\mathbf{c}} \mathbf{H}_{Y=3|X^*,\mathbf{c}} \mathbf{H}_{X^*,Z|\mathbf{c}} \tag{4.8}$$

$$\mathbf{H}_{X,Z|\mathbf{c}} = \mathbf{H}_{X|X^*,\mathbf{c}} \mathbf{H}_{X^*,Z|\mathbf{c}} \tag{4.9}$$

$$\mathbf{H}_{Y=3,X,Z|\mathbf{c}} {\mathbf{H}_{X,Z|\mathbf{c}}}^{-1} = \mathbf{H}_{X|X^*,\mathbf{c}} \mathbf{H}_{Y=3|X^*,\mathbf{c}} {\mathbf{H}_{X|X^*,\mathbf{c}}}^{-1} \tag{4.10}$$

[5]Theoretically, the three measurements can be used interchangeably. Empirically, we choose FCC to be the third measurement of SWB as it serves the purpose of measuring SWB, but it also causes variation in the joint distribution of X and Z by each level of Z to the extent that $\mathbf{H}_{X,Z|\mathbf{c}}$ has the best chance to exist. Namely, if X and Z are identical, Z does not provide any additional information on their joint distribution; and each column of $\mathbf{H}_{X,Z|\mathbf{c}}$ is expressed in the other two columns, then the inverse of $\mathbf{H}_{X,Z|\mathbf{c}}$ will not exist.

The last equation indicates that the left-hand side has an eigenvalue-eigenvector decomposition on the right hand side. The left hand side is obtained from the survey responses. Accordingly, the misclassification matrix $\mathbf{H}_{X|X^*,\mathbf{c}}$ is directly estimated through eigen-decomposition.

Theorem 1. *Under Assumptions 1 through 5, the misclassification matrix* $\mathbf{H}_{X|X^*,\mathbf{c}}$ *is nonparametrically identified and directly estimated through the unique eigenvalue-eigenvector decomposition of equation 10.*

This result is obtained following Theorem 1 in Hu (2008); See a complete proof in Section S2 of the Supplement. In the second step, the distribution of the underlying true status is recovered by multiplying the inverse matrix of $\mathbf{H}_{X|X^*,\mathbf{c}}$ by the self-reported SWB vector. Reversibility is guaranteed by the property of eigen-vectors. Since self-reported status is directly observed in the survey data, the distribution vector of the true latent SWB is obtained by:

$$\left\{\begin{array}{c} Pr(X^*=1|\mathbf{c}) \\ Pr(X^*=2|\mathbf{c}) \\ Pr(X^*=3|\mathbf{c}) \end{array}\right\} = \mathbf{H}_{X|X^*,\mathbf{c}}{}^{-1} \left\{\begin{array}{c} Pr(X=1|\mathbf{c}) \\ Pr(X=2|\mathbf{c}) \\ Pr(X=3|\mathbf{c}) \end{array}\right\} \tag{4.11}$$

4.4 Results

4.4.1 Reported versus Corrected Happiness

In total, our sample covers 80 countries around the world.[6] Figure 4.1 presents a map of the average reported (Panel a) and corrected (Panel b) happiness for all 80 countries. A larger numerical value represents a higher level of happiness.

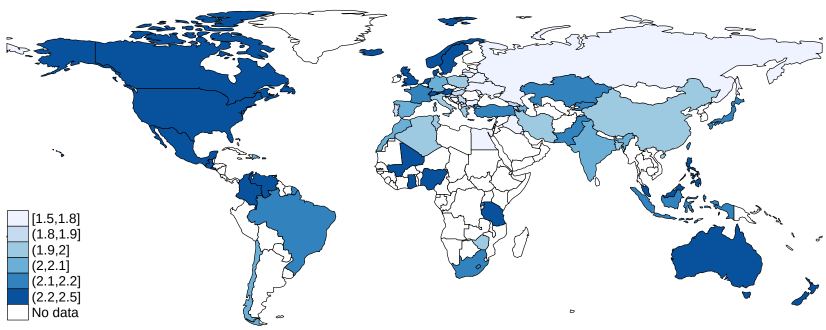
[6] The list of countries is provided in Table S2 of the supplementary material. There are at most 18 demographic groups in each country, i.e., full factorial of two genders, three age groups, and three income groups. Due to sample size limitations for certain countries, not all demographic groups within those countries were identified. Following the assumptions of Section 2.3, over 18 demographic groups, 24 countries are fully identified, 18 countries are identified for 17 demographic groups, 25 countries are identified for 16 demographic groups, 9 countries are identified for 15 demographic groups and 2 countries are identified for 14 demographic groups. Binomial tests for the hypothesis that each identified group is chosen with equal probability (1/18) can not be rejected for all 18 demographic subsamples; see the supplementary material for more details.

This variable is obtained as the summation of happiness measures weighted by the proportion of reported or corrected happiness at each level. Specifically, the observed proportion of people in country i who report being "unhappy", "quite happy" and "very happy" are defined by the vector $\{\omega_{1i}, \omega_{2i}, \omega_{3i}\}$. Using the constructive identification procedures described in the previous section, the underlying latent distribution for the three levels of SWB is represented by the vector $\{\omega^*_{1i}, \omega^*_{2i}, \omega^*_{3i}\}$.

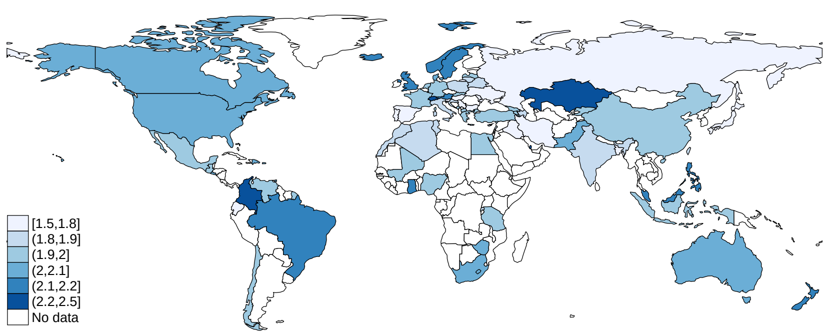
$$\text{Reported Happiness} = \quad 1 * \omega_{1i} + 2 * \omega_{2i} + 3 * \omega_{3i} \tag{4.12}$$

$$\text{Corrected Happiness} = \quad 1 * \omega^*_{1i} + 2 * \omega^*_{2i} + 3 * \omega^*_{3i} \tag{4.13}$$

By comparing panels (a) and (b), we find that while nations such as Russia, China and Brazil do not appear to have sizable differences between reported and corrected levels of happiness, other countries show pronounced differences (e.g., Japan, United States).



(a) Reported Happiness



(b) Corrected Happiness

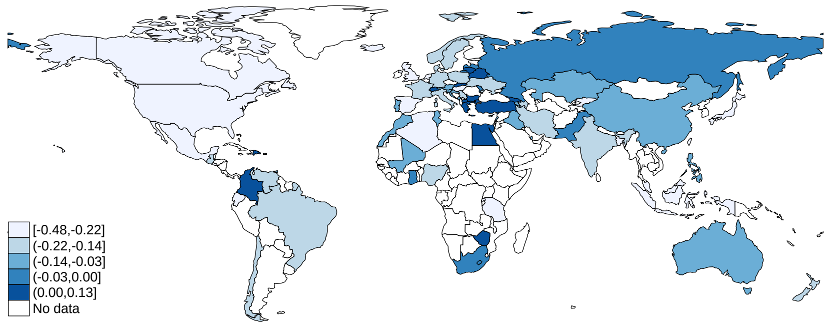
Figure 4.1: World Map: Reported vs. Corrected Happiness

The difference between reported and corrected SWB at each level of happiness can be either positive or negative, (i.e., over-reporting or under-reporting the true happiness status).[7] Simply comparing the aggregate levels of SWB fails to account for the direction of the misclassification errors. In order to further explore discrepancies between reported and corrected happiness, we also calculate misclassification errors taking into account the direction of correction and the distance between the two vectors. Specifically, we utilize the potential error at each level of happiness to construct misclassification errors of SWB for each country. We define **Level 1** bias as $\omega_{1i} - \omega_{1i}^{*}$ and **Level 3** bias as $\omega_{3i} - \omega_{3i}^{*}$, which represent the biases in reporting "unhappy" and "very happy" categories, respectively. These two measures are further examined since they are more informative about the direction of misclassification errors of individual SWB. A positive **Level 3** (**Level 1**) bias implies a significant proportion of people reporting more extreme feelings of happiness than they really are; i.e., reporting "very happy" ("unhappy"), while actually being "quite happy". In contrast, a negative **Level 3** (**Level 1**) bias suggests that people are inclined to suppress extreme categories. For instance, an individual may overstate happiness due to self-deception, while others may understate SWB. In another example, as the "Doctrine of the Mean" is highly praised in Asian Confucian culture,[8] citizens in Confucian countries may be more likely to report being "quite happy" and avoid extreme categories.

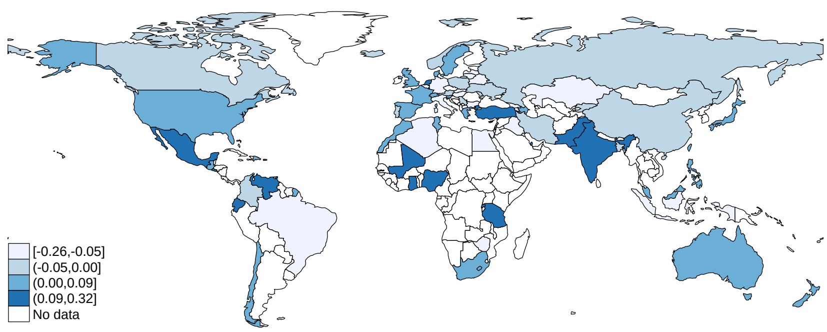
Figure 4.2 summarizes **Level 1** and **Level 3** biases across countries. Many countries have substantial bias at both levels. Misclassification errors in countries such as the United States and Canada come mainly from **Level 1** bias. E.g., American and Canadian respondents tend to avoid reporting that they are "unhappy". It is also observed that China, Hong Kong and Germany have a negative bias at both levels simultaneously. Note that countries that have negative biases at both levels are not only restricted to confucian or Asian countries. Also, some countries present positive

[7]We use "over-report", hereafter, if reported SWB (by level) is larger than corrected SWB (by level); and we use "under-report", hereafter, if reported SWB (by level) is smaller than corrected SWB (by level). We define "upward (downward) shifting", hereafter, in situations where the SWB status is reported upward (downward) from the underlying true status of SWB.

[8]As interpreted by James Legge in The Sacred Books of China, the principle of the "Doctrine of the Mean" is to guide the mind to a state of constant equilibrium, which recommends one should never act in excess. James Legge was a Scottish sinologist, missionary, and scholar, best known as an early translator of Classical Chinese texts into English.



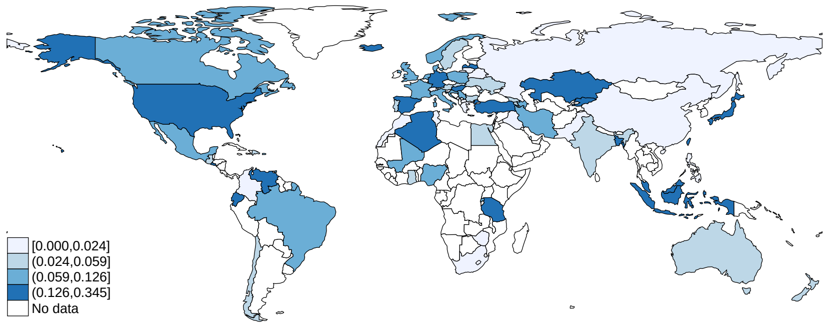
(a) **Level 1** Bias (Unhappy)



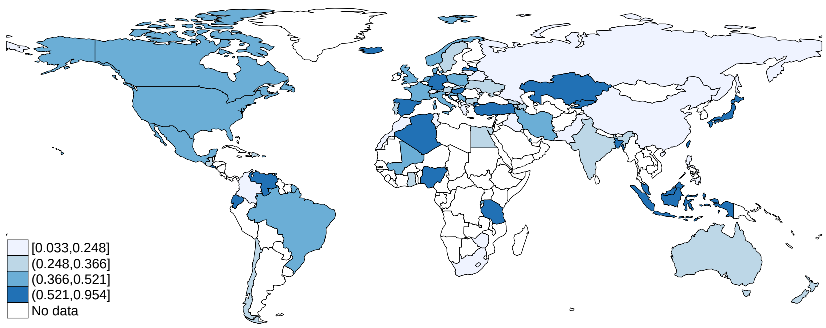
(b) **Level 3** Bias (Very Happy)

Figure 4.2: World Map: Misclassification Error of Happiness.A

Level 1 bias but negative **Level 3** bias; which means they are more inclined to over-report being "unhappy", but under-report being "very happy". There are nine countries with this pattern, six of which are from eastern Europe, including Belarus, Bulgaria, Georgia, Lithuania, Slovakia and Macedonia. The other three countries are Colombia, Egypt, and Zimbabwe. On the other hand, there are 38 countries with negative **Level 1** and positive **Level 3** biases, implying under-reporting being "unhappy" and over-reporting being "very happy". Out of the 38 countries sharing this pattern, 13 countries have both **Level 1** and **Level 3** biases larger than 0.1 in absolute value. These countries are Cyprus, Ecuador, El Salvador, India, Mexico, Netherlands, Nigeria, Puerto Rico, Qatar, Rwanda, Singapore, Tanzania and Venezuela, most of which are located in tropical regions. As noticed, misclassification error patterns are not specific to certain climate zones or geographical regions. Hence, we explore to what degree other cultural factors such as major religious beliefs and demographic composition also play a role in the direction and magnitude of **Level 1** and **Level 3** biases.



(a) ME-Euclidean



(b) ME-rectilinear

Figure 4.3: World Map: Misclassification Error of Happiness.B

We represent misclassification errors in vector distance using Euclidean distance and rectilinear distance (L1 norm). The misclassification error vectors measure the distance between *reported* $\{\omega_{1i}, \omega_{2i}, \omega_{3i}\}$ and *corrected* happiness $\{\omega^*_{1i}, \omega^*_{2i}, \omega^*_{3i}\}$. Specifically,

$$\text{Euclidean distance} = \sum_{j=1}^{3}(\omega_{ji} - \omega^*_{ji})^2 \tag{4.14}$$

$$\text{Rectilinear distance} = \sum_{j=1}^{3}|\omega_{ji} - \omega^*_{ji}| \tag{4.15}$$

Figure 4.3 maps the two measurements of misclassification for Euclidean distance (panel a) and rectilinear distance (panel b). By integrating the three happiness levels, Euclidean and rectilinear measures display similar orderings over the 80 countries. As seen in Figure 4.3, in both cases the countries with the smallest misclassification errors are Colombia, Russia and Moldova. The countries with the largest errors (in both measurements) are Ecuador, Hong Kong and Indonesia. For all other countries, the detailed orderings are slightly different for the two measures. For the three most populous countries in the world, both measures indicate that China has relatively small misclassification errors, while the United States experiences sizable misclassification errors, and India is subject to moderate errors.[9]

4.4.2 Determinants of Misclassification Errors in SWB

One of the main objectives of this article is to understand why some countries display substantial misclassification errors of SWB while other countries do not. The next step after uncovering the misclassification errors of SWB is to take a closer look at the relationship between country or individual level characteristics and the magnitude of misclassification errors of SWB.

4.4.2.1 Correlation with Country-level Characteristics

We start by conducting a regression on the association between misclassification errors and country-level characteristics. Then, we examine the correlation between misclassification errors and demographic characteristics while controlling for country fixed effects.

The results from regressing various measures of misclassification errors on the log of GDP *per capita* and religious beliefs are shown in Table 4.1. The first row suggests that none of the measurements of misclassification errors are correlated with the log of GDP *per capita*. In the following rows of Table 4.1, we examine the correlation between misclassification errors and prevalent religious beliefs. In columns 2 and 4, the outcome variables are misclassification errors measured in vector Euclidean and rectilinear distance. Confucian countries show larger errors compared to other countries.

In order to explore the direction of bias, columns 6 and 8 implement analogous estimations with the outcome variables focusing on **Level 1** and **Level 3** biases of the reported SWB. Point estimates for **Level 1** bias suggest that compared to predominantly Christian countries, which are the majority in our sample, “unhappy" people in Confucian countries tend to report being “quite happy". With respect to **Level 3** bias, only the mixed religion category shows a statistically significant effect. That is, for countries with multiple dominant religions, a larger proportion of

[9]We present the estimated misclassification errors across all 80 countries in Table S2 of the supplemental material.

Table 4.1: Determinants of Misclassification Error of SWB

	(1) Euclidean	(2) Euclidean	(3) Rectilinear	(4) Rectilinear	(5) Level 1	(6) Level 1	(7) Level 3	(8) Level 3
log of GDP per capita	0.004		0.012		-0.013		-0.017	
	(0.008)		(0.021)		(0.013)		(0.013)	
Christianity		-		-		-		-
		(.)		(.)		(.)		(.)
Confucian		0.086**		0.179**		-0.103*		0.041
		(0.033)		(0.087)		(0.056)		(0.056)
Hinduism		-0.036		-0.078		-0.014		0.088
		(0.071)		(0.187)		(0.120)		(0.121)
Mixed		0.001		0.033		0.056		0.120*
		(0.042)		(0.110)		(0.071)		(0.071)
Islam		0.023		0.062		-0.006		0.014
		(0.019)		(0.050)		(0.032)		(0.032)
Constant	0.069***	0.069***	0.361***	0.366***	-0.104***	-0.130***	0.066*	0.015
	(0.022)	(0.010)	(0.056)	(0.026)	(0.036)	(0.017)	(0.036)	(0.017)
Observations	78	79	78	79	78	79	78	79

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Mixed indicates countries with more than one major religion.

Andorra and Puerto Rico excluded due to the lack of GDP data for these two countries.

Level 1 indicates the difference between reported and corrected value in level of "Unhappy".

Level 3 indicates the difference between reported and corrected value in level of "Very Happy".

people who are "quite happy" report being "very happy".

4.4.2.2 Correlation with Demographic Characteristics

In order to investigate the contributors of misclassification error of SWB in a more disaggregated way, we pooled data from all countries and regressed misclassification errors on individual level characteristics (i.e., income, gender, age). The results are shown in Table 4.2.

Gender differences in misclassification errors of SWB seem to be negligible. Male respondents have slightly larger misclassification errors measured by rectilinear distance. We do not observe significant variations in misclassification errors by age measured by Euclidean or rectilinear distance. However, in columns 3 and 4, we observe that younger individuals are more inclined to over-report being "very happy" and under-report being "unhappy", compared to older individuals. With respect to individual income, we find that low-income individuals are more likely to report being "quite happy" when they are actually "unhappy"; and they tend to report being "very happy" when they are in fact "quite happy". This finding indicates an overall pattern of upward shifting from the true underlying SWB when low-income individuals respond to happiness surveys. Interestingly, relative to middle-income individuals, the high-income group shows a similar pattern in **Level 3** bias to the low-income group, but with smaller magnitude and lower significance. In order to capture heterogenous correlations of misclassification errors of SWB and individual-level characteristics conditional on country-level features, we split the sample by prevalent religion and economic development stage. The results are reported in Tables 4.3 and 4.4 respectively.

Recall that in Table 4.1, we show that Confucianism affects the aggregate country level misclassification errors. It is also possible that religion, in general, may also affect how misclassification errors of SWB relate to gender, age and individual income. In columns 1 and 2 of Table 4.3, males have a positive but small effect on misclassification errors of SWB in Confucian countries, while in Islamic and Christian countries males are not significantly different from females. On the one hand, with respect to **Level 1** and **Level 3** biases, "unhappy" men from Christian countries tend to under-report their true underlying SWB. On the other hand, "very happy" men from Islamic countries are more likely to report being "quite happy" or "unhappy". We also find heterogeneous effects by age. In line with the results for the overall sample, in predominantly Christian countries, younger people tend to express optimistic feelings when reporting SWB, generally over-rating their subjective well-being compared to older cohorts.

Young people from Islamic countries show a similar pattern. However, people from Confucian countries show an opposite pattern with regards to the association between age and misclassification errors of SWB. That is, they tend to shift their overall underlying SWB downwards.

Regarding relative income, people from countries with multiple prevalent religions show a similar pattern. In general, low-income and high-income cohorts are more likely to report a higher overall subjective well-being level than their true underlying status.

We further probe the correlation between misclassification errors of SWB and individual characteristics conditional on the development stage of the country. Table 4.4 provides suggestive evidence of heterogenous relationships across countries in different stages of development. Gender has no effect on misclassification errors of SWB in low and middle-income countries, but

Table 4.2: Misclassification Error of SWB, By Group

	(1) ME-Euclidean	(2) ME-rectilinear	(3) **Level 1**	(4) **Level 3**
Male	0.024	0.039*	-0.022	-0.016
	(0.016)	(0.023)	(0.015)	(0.014)
15-29 years	0.015	0.022	-0.037**	0.055***
	(0.020)	(0.028)	(0.018)	(0.017)
30-49 years	0.005	0.013	-0.015	0.010
	(0.020)	(0.028)	(0.018)	(0.017)
Income Level: Low	0.038*	0.053*	-0.082***	0.072***
	(0.020)	(0.029)	(0.018)	(0.017)
Income Level: High	0.043**	0.068**	-0.092***	0.032*
	(0.020)	(0.028)	(0.018)	(0.017)
Constant	0.248***	0.644***	-0.045**	-0.017
	(0.020)	(0.028)	(0.018)	(0.017)
Observations	1301	1301	1301	1301

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Level 1 indicates the difference between reported and corrected value in level of “Unhappy”. **Level 3** indicates the difference between reported and corrected value in level of “Very Happy”. The reference group for age band is “50 and more years”; The reference group for income is “middle income level”. All specifications controlled for country level fixed effects.

males from high-income countries show higher misclassification errors in SWB measured by Euclidean and rectilinear distance. In particular, in high-income countries, males tend to report being “quite happy” or “very happy” when they are actually “unhappy”. Young people from low-income countries tend to report a “very happy” status when their true underlying happiness status is lower. This bias is relatively weak in middle-income countries and it is not significant in high-income countries. Low-income individuals tend to report their true underlying SWB in an upward-biased pattern. Moreover, the high-income group shares the same reporting pattern of SWB as the low-income group. The coefficients are significant for both cohorts in low-income countries. In middle-income countries, both low-income and high-income people who are actually “unhappy” tend to report they are “quite happy” or “very happy”. Additionally, low and high-income people living in high-income countries present larger misclassification errors measured by Euclidean and rectilinear distance.

Table 4.3: Misclassification Error of SWB, Different Religions

	(1) ME-Euclidean	(2) ME-rectilinear	(3) **Level 1**	(4) **Level 3**
		Confucian Countries		
Male	0.109	0.155*	-0.089	-0.079
	(0.072)	(0.093)	(0.061)	(0.048)
15-29 years	0.016	0.052	0.139*	-0.121**
	(0.090)	(0.115)	(0.076)	(0.059)
30-49 years	-0.066	-0.082	0.159**	0.002
	(0.087)	(0.112)	(0.074)	(0.057)
Income Level: Low	0.035	0.051	-0.133*	0.173***
	(0.089)	(0.115)	(0.076)	(0.059)
Income Level: High	0.009	-0.016	-0.104	0.120**
	(0.088)	(0.113)	(0.075)	(0.058)
Observations	81	81	81	81
		Christian Countries		
Male	0.024	0.039	-0.030*	0.008
	(0.020)	(0.029)	(0.018)	(0.016)
15-29 years	-0.001	-0.002	-0.037*	0.051**
	(0.024)	(0.035)	(0.022)	(0.020)
30-49 years	0.029	0.054	-0.045**	0.012
	(0.024)	(0.035)	(0.022)	(0.020)
Income Level: Low	0.046*	0.063*	-0.098***	0.041**
	(0.025)	(0.036)	(0.022)	(0.020)
Income Level: High	0.059**	0.086**	-0.098***	0.023
	(0.024)	(0.035)	(0.022)	(0.020)
Observations	838	838	838	838
		Islamic Countries		
Male	-0.002	-0.003	0.028	-0.062**
	(0.036)	(0.048)	(0.030)	(0.031)
15-29 years	0.037	0.044	-0.052	0.082**
	(0.044)	(0.059)	(0.037)	(0.038)
30-49 years	-0.010	-0.024	0.023	-0.000
	(0.044)	(0.059)	(0.037)	(0.038)
Income Level: Low	-0.014	-0.027	-0.040	0.122***
	(0.044)	(0.060)	(0.037)	(0.038)
Income Level: High	0.007	0.039	-0.080**	0.041
	(0.043)	(0.059)	(0.037)	(0.037)
Observations	315	315	315	315

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All specifications controlled for country level fixed effects.

Table 4.4: Misclassification Error of SWB, Different Development Stages

	(1) ME-Euclidean	(2) ME-rectilinear	(3) **Level 1**	(4) **Level 3**
		Low Income Countries		
Male	-0.051	-0.064	0.015	-0.031
	(0.033)	(0.045)	(0.030)	(0.030)
15-29 years	0.035	0.057	-0.054	0.109***
	(0.041)	(0.055)	(0.037)	(0.037)
30-49 years	0.014	0.033	0.006	0.040
	(0.040)	(0.055)	(0.036)	(0.037)
Income Level: Low	0.069*	0.076	-0.103***	0.167***
	(0.041)	(0.055)	(0.037)	(0.037)
Income Level: High	0.034	0.059	-0.142***	0.112***
	(0.040)	(0.054)	(0.036)	(0.037)
Observations	334	334	334	334
		Middle Income Countries		
Male	0.023	0.057	-0.010	0.011
	(0.029)	(0.041)	(0.027)	(0.025)
15-29 years	-0.026	-0.039	-0.029	0.051*
	(0.036)	(0.051)	(0.033)	(0.031)
30-49 years	-0.016	-0.008	0.023	0.012
	(0.037)	(0.051)	(0.034)	(0.031)
Income Level: Low	-0.024	-0.035	-0.069**	0.031
	(0.036)	(0.051)	(0.034)	(0.031)
Income Level: High	-0.001	0.005	-0.067**	-0.020
	(0.036)	(0.050)	(0.033)	(0.030)
Observations	386	386	386	386
		High Income Countries		
Male	0.067***	0.086**	-0.049**	-0.024
	(0.024)	(0.036)	(0.022)	(0.019)
15-29 years	0.031	0.043	-0.031	0.026
	(0.030)	(0.044)	(0.026)	(0.023)
30-49 years	0.012	0.013	-0.051*	-0.008
	(0.030)	(0.043)	(0.026)	(0.023)
Income Level: Low	0.060**	0.096**	-0.078***	0.044*
	(0.030)	(0.044)	(0.026)	(0.023)
Income Level: High	0.079***	0.117***	-0.080***	0.023
	(0.030)	(0.043)	(0.026)	(0.023)
Observations	581	581	581	581

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All specifications controlled for country level fixed effects.

Countries ranked as lower-middle-income are integrated with low-income countries.

4.4.3 The Easterlin Paradox

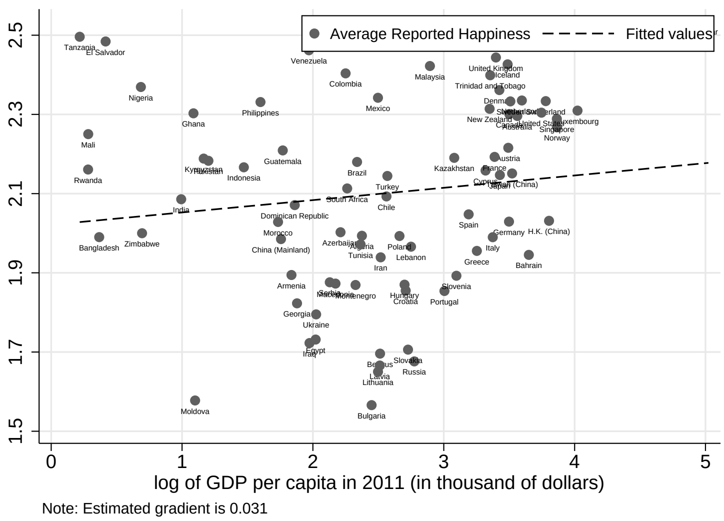
In this section, we revisit the Easterlin paradox and modified Easterlin Hypothesis by investigating the correlation between reported and corrected happiness and GDP *per capita*. We first implement simple regressions of average reported and corrected happiness on the log of GDP *per capita*. As shown in Table 4.5, there is no significant correlation between reported happiness and GDP *per capita*. This result, based on the reported level of happiness, provides support for the Easterlin paradox. However, the corrected measure of happiness is strongly associated with GDP *per capita*, leading to reject the Easterlin hypothesis. Notably, the coefficients are only slightly different.

Table 4.5: Easterlin Paradox, Simple Regressions

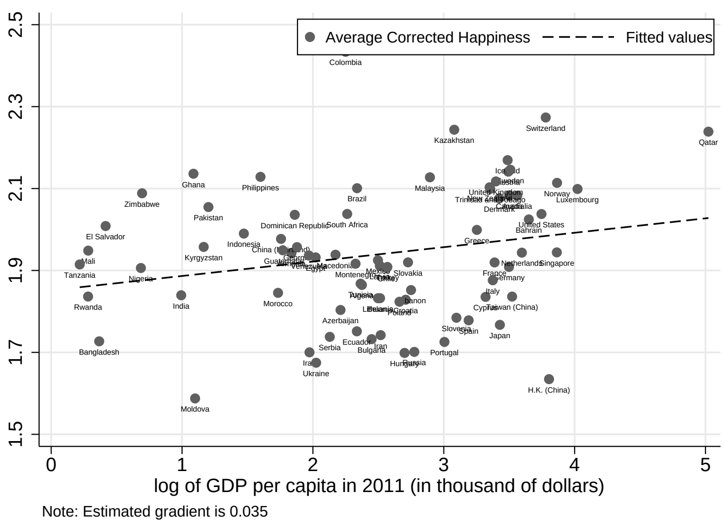
	(1) Average Reported Happiness	(2) Average Corrected Happiness
log of GDP per capita	0.031	0.035**
	(0.027)	(0.018)
Constant	2.021***	1.851***
	(0.073)	(0.047)
adj.R-squared	0.004	0.038
Observations	78	78

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Andorra and Puerto Rico are excluded due to the lack of GDP data for these two countries.
The log of *per capita* GDP is adjusted in 2011 in thousand of dollars.

From the scatter diagram ("Fig 1" in the original paper) of the seminal paper by Easterlin (1974), the author indicates that the "the association between the wealth and happiness ... is not so clear-cut". We replicate Easterlin's findings using reported happiness from WVS/EVS and the log of GDP *per capita* instead of GNP *per capita*. The equivalent representations are displayed in the two panels of the scatter plots in Figure 4.4. Similar to the original graph, panel (a) of Figure 4.4 does not show a definitive association between SWB and wealth. However, in panel (b) this relationship is statistically significant when *corrected* happiness is used to measure SWB.



(a) Reported Happiness and GDP per capita



(b) Corrected Happiness and GDP per capita

Figure 4.4: Cross-Country Analysis of Happiness and GDP per capita

Only a minor change in the slope is observed when comparing the two fitted lines for *reported* and *corrected* happiness.

We also test the modified-Easterlin hypothesis following Stevenson and Wolfers (2013):

$$\text{Subjective Well-being}_c = \alpha + \beta_{poor} * I(GDP_c < k) \times [log(GDP_c) - log(k)] + \beta_{rich} * I(GDP_c \geq k) \times [log(GDP_c) - log(k)] + \epsilon_c$$

where GDP_c stands for GDP per capita of country C, and k is the threshold of satiation point (e.g., \$15,000 in Stevenson and Wolfers (2013)).

If the modified-Easterlin hypothesis holds, the equality of $\beta_{rich} = 0$ must hold, or at least the relationship of $\beta_{rich} < \beta_{poor}$. However, we find strong evidence against the modified-Easterlin hypothesis, as shown in Table 4.6. In columns 1 and 2 of panel A with \$10,000 as the threshold, we find that $\beta_{rich} > 0 \geq \beta_{poor}$, regardless of whether the dependent variable is reported happiness or corrected happiness. The reported and corrected happiness measures are positively associated with the log of GDP *per capita* for high-income countries at the 1% level of significance. Notice that using the reported data, the estimation yields a negative SWB-GDP gradient among poor countries, and this slope coefficient is also significant as shown in column 1. However, this negative coefficient becomes insignificant with the corrected data, suggesting that the SWB-GDP relation is not significantly different from zero for the countries below the threshold after correcting the misclassification errors. We further test whether this relationship is sensitive to the selected threshold amount and the exact specification. The result of $\beta_{rich} > \beta_{poor}$ is robust to the variation of the threshold to \$15,000 and \$20,000. These results are consistent with the findings in Stevenson and Wolfers (2013).

It is likely that there are heterogenous correlations between GDP and SWB across countries in different development stages. Hence, we also specify a flexible nonparametric linear spline. The results are presented in panel b of Table 4.6 with knots set at \$10,000 and \$20,000. The point estimate is similar to that in panel a and significant at the 1% level. Although there is a reduction

Table 4.6: Modified Easterlin Paradox

	(1) Reported	(2) Corrected	(3) Reported	(4) Corrected	(5) Reported	(6) Corrected	(7) Reported	(8) Corrected
Panel a								
rich10=0 × lgdp10	-0.165***	-0.031						
	(0.047)	(0.034)						
rich10=1 × lgdp10	0.220***	0.099***						
	(0.046)	(0.033)						
rich15=0 × lgdp15			-0.126***	-0.027				
			(0.038)	(0.027)				
rich15=1 × lgdp15			0.319***	0.149***				
			(0.060)	(0.043)				
rich20=0 × lgdp20					-0.086**	-0.014		
					(0.035)	(0.025)		
rich20=1 × lgdp20					0.391***	0.186***		
					(0.083)	(0.058)		
Panel b								
$< lgdp10$							-0.149***	-0.011
							(0.051)	(0.036)
$> lgdp10, < lgdp20$							0.117	-0.024
							(0.126)	(0.090)
$> lgdp20$							0.297***	0.191***
							(0.099)	(0.070)
Constant	1.930***	1.878***	1.934***	1.879***	1.960***	1.892***	2.296***	1.931***
	(0.042)	(0.030)	(0.041)	(0.030)	(0.044)	(0.031)	(0.090)	(0.064)
adj.R-squared	0.229	0.086	0.258	0.121	0.207	0.112	0.226	0.101
Observations	78	78	78	78	78	78	78	78

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

rich10=1 indicates countries with GDP over $10,000 per capita, similar definition for other thresholds.

lgdp15 is the corresponding log difference over this threshold as in Equation 4.16. The similar notations applies for lgdp10 and lgdp20.

in the magnitude of the correlation between the log of GDP *per capita* and SWB after applying the SWB correction in all specifications, it is clear that this association is robust even beyond the threshold of $20,000. Taken together, neither the original nor the modified-Easterlin hypothesis is supported by the cross-country analysis of corrected SWB. This means that *corrected* measures of SWB are correlated with GDP *per capita*; However, *reported* SWB is not. This implies that SWB measures can be used to evaluate economic development and welfare, but they have to be corrected for measurement errors.

4.5 Conclusion

Evaluations of economic development and social welfare normally use objective outcome indicators; however, subjective measurements of welfare also provide insights. Subjective well-being data have usually been obtained by self-reported surveys. However, due to its inherent nature, reported SWB is subject to substantial misclassification errors.

In this article we apply a new method to uncover the latent true distribution of subjective well-being using data from the integrated EVS/WVS surveys. We find that reported happiness in surveys has substantial misclassification errors. This is of particular importance to the literature since most studies rely on reported SWB to address economic questions of interest. We explore the characteristics of countries and demographic groups with substantial misclassification errors and find that religious beliefs and the development stage of the countries play a critical role in the magnitude of misclassification errors in reported SWB. We revisit the Easterlin paradox and find that based on a country level analysis there is no evidence supporting neither the original hypothesis nor the modified version of the Easterlin's paradox when using the *corrected* measure of happiness, although the *reported* (uncorrected) SWB provides support to the original Easterlin paradox.

The identified misclassification errors in reported subjective well-being provide directions for the future research. First, further efforts can be devoted to conduct causal analysis of the relationship between misclassification errors and their influencing factors. A richer longitudinal dataset or exogenous variations can help to establish causality. In addition, follow-up surveys may enable researchers to explore individual characteristics driving misclassification errors in SWB reports.

Finally, experimental studies can be useful to examine how psychological and behavioral elements such as envy and self-deception affect misclassification errors of SWB measures in survey data.

4.6 Supplement

4.6.1 Data Summary Statistics

Data analyzed for this study is from the Integrated European Survey (EVS) and World Value Survey (WVS) from 1981 to 2014. During these years, EVS was conducted for four waves: 1981-1984, 1990-1993, 1999-2001 and 2008-2010, while WVS was conducted for six waves: 1981-1984, 1989-1993, 1994-1998, 1999-2004, 2005-2009, 2010-2014.

Figure 4.5 provides a list of the countries with respect to their corresponding survey waves. Some countries were not continuously surveyed across the waves. For the same wave, WVS and EVS may or may not cover exactly the same years (e.g., "WVS: 1989-1993" vs. "EVS: 1990-1993", "WVS: 1999-2004" vs. "EVS: 1999-2001"). Due to this very unbalanced feature of the data structure, we pool all the waves together for each country and conduct a cross sectional analysis.

Table 4.7 shows the summary statistics for the reported values of the three measurements of subjective well-being (SWB). In each demographic cell, it reports the sample mean for the percentage of people reporting each level of a given SWB measure. The tabulation is implemented with all 80 countries sample pooled together.

In Table 4.8, we present the estimated misclassification error in each country. It summarizes by major religion, geographical location, development stage, reported/corrected SWB along with the misclassification errors (by level and vector distance).

Figure 4.5: Survey Waves for the 80 Countries in the Sample

Table 4.7: Summary Statistics (%)

Panel A: Happiness			
	Not Happy	Quite Happy	Very Happy
Income Level: Low	27.64	49.93	22.43
Income Level: Middle	16.50	55.62	27.88
Income Level: High	10.19	53.13	36.68
15-29 years	15.79	53.32	30.90
30-49 years	18.42	54.71	26.87
50 and more years	22.21	53.76	24.03
Male	18.80	54.81	26.40
Female	18.94	53.53	27.53
Panel B: Life Satisfaction (LS)			
	Most Dissatisfied	Quite Satisfied	Most Satisfied
Income Level: Low	27.87	52.89	19.23
Income Level: Middle	14.87	62.82	22.31
Income Level: High	9.38	58.39	32.23
15-29 years	15.80	59.66	24.54
30-49 years	18.22	59.51	22.27
50 and more years	18.80	56.75	24.45
Male	17.93	59.27	22.80
Female	17.66	58.18	24.16
Panel C: Freedom of Choice and Control in Life (FCC)			
	Not At All	Quite A Bit	A Great Deal
Income Level: Low	23.06	52.93	24.02
Income Level: Middle	13.55	62.39	24.05
Income Level: High	9.83	59.36	30.81
15-29 years	14.21	59.27	26.52
30-49 years	15.70	59.88	24.42
50 and more years	17.62	57.77	24.61
Male	15.09	58.67	26.24
Female	16.72	59.38	23.90

Notes: sample means in percentage for 80 countries in the sample

Table 4.8: Misclassification Error at Country Level

Country	**Religion**	**Area**	**Development**	*Reported*	*Corrected*	**Misclassification Error**			
						Euclidean	*Rectilinear*	*Level 1*	*Level 3*
Algeria	Islam	GME	Upper middle	1.994	1.865	0.158	0.632	-0.223	-0.094
Andorra	Christianity	Europe	High	2.206	1.633	0.419	0.985	-0.493	0.081
Armenia	Christianity	Europe	Lower middle	1.894	1.943	0.035	0.301	-0.051	-0.099
Australia	Christianity	Oceania	High	2.296	2.083	0.026	0.261	-0.131	0.083
Austria	Christianity	Europe	High	2.216	2.141	0.034	0.290	-0.110	-0.035
Azerbaijan	Islam	Europe	Upper middle	2.003	1.804	0.059	0.361	-0.180	0.019
Bahrain	Islam	GME	High	1.945	2.025	0.040	0.314	-0.039	-0.118
Bangladesh	Islam	Asia	Lower middle	1.990	1.727	0.146	0.544	-0.268	-0.005
Belarus	Mixed	Europe	Upper middle	1.696	1.832	0.014	0.194	0.040	-0.097
Brazil	Christianity	Americas	Upper middle	2.180	2.101	0.073	0.433	-0.148	-0.069
Bulgaria	Christianity	Europe	Upper middle	1.566	1.732	0.029	0.267	0.033	-0.134
Canada	Christianity	Americas	High	2.301	2.083	0.100	0.450	-0.221	-0.004
Chile	Christianity	Americas	High	2.093	1.906	0.058	0.352	-0.176	0.011
P.R.China	Confucian	Asia	Upper middle	1.985	1.977	0.012	0.181	-0.049	-0.041
Colombia	Christianity	Americas	Upper middle	2.404	2.434	0.000	0.033	0.017	-0.013
Croatia	Christianity	Europe	High	1.855	1.829	0.066	0.418	-0.118	-0.091
Cyprus	Christianity	Europe	High	2.158	1.835	0.054	0.362	-0.142	0.181
Denmark	Christianity	Europe	High	2.362	2.076	0.075	0.436	-0.218	0.068

Continued on next page

Table 4.8 Continued

Country	Religion	Area	Development	*Reported*	*Corrected*	Misclassification Error			
						Euclidean	*Rectilinear*	*Level 1*	*Level 3*
Dominican Rep.	Christianity	Americas	Upper middle	2.071	2.036	0.031	0.285	0.054	0.089
Ecuador	Christianity	Americas	Upper middle	2.512	1.751	0.345	0.954	-0.477	0.284
Egypt	Islam	GME	Lower middle	1.732	1.932	0.024	0.250	0.125	-0.075
El Salvador	Christianity	Americas	Lower middle	2.484	2.008	0.133	0.592	-0.296	0.180
France	Christianity	Europe	High	2.193	1.920	0.074	0.430	-0.215	0.058
Georgia	Christianity	Europe	Upper middle	1.823	1.957	0.032	0.258	0.005	-0.129
Germany	Christianity	Europe	High	2.029	1.909	0.157	0.631	-0.218	-0.097
Ghana	Christianity	SSA	Lower middle	2.303	2.136	0.055	0.331	-0.001	0.166
Great Britain	Christianity	Europe	High	2.444	2.117	0.105	0.512	-0.256	0.070
Greece	Christianity	Europe	High	1.955	1.999	0.009	0.146	0.058	0.015
Guatemala	Christianity	Americas	Lower middle	2.209	1.949	0.047	0.353	-0.176	0.083
Hong Kong(China)	Confucian	Asia	High	2.031	1.635	0.393	0.916	-0.427	-0.031
Hungary	Christianity	Europe	High	1.870	1.699	0.175	0.655	-0.249	-0.078
Iceland	Christianity	Europe	High	2.426	2.169	0.165	0.593	-0.277	-0.020
India	Hinduism	Asia	Lower middle	2.086	1.839	0.033	0.288	-0.144	0.102
Indonesia	Islam	Asia	Lower middle	2.166	1.990	0.199	0.699	-0.263	-0.086
Iran	Islam	GME	Upper middle	1.939	1.742	0.082	0.410	-0.201	-0.004
Iraq	Islam	GME	Upper middle	1.723	1.700	0.023	0.246	-0.073	-0.050
Italy	Christianity	Europe	High	1.990	1.876	0.099	0.497	-0.181	-0.067

Continued on next page

Table 4.8 Continued

Country	**Religion**	**Area**	**Development**	*Reported*	*Corrected*	**Misclassification Error**			
						Euclidean	*Rectilinear*	*Level 1*	*Level 3*
Japan	Confucian	Asia	High	2.147	1.767	0.134	0.583	-0.291	0.088
Kazakhstan	Islam	Asia	Upper middle	2.190	2.244	0.159	0.647	-0.135	-0.188
Kyrgyzstan	Islam	Asia	Lower middle	2.188	1.957	0.177	0.633	-0.274	-0.043
Latvia	Christianity	Europe	High	1.666	1.912	0.136	0.531	-0.010	-0.255
Lebanon	Islam	GME	Upper middle	1.966	1.852	0.013	0.181	-0.091	0.023
Lithuania	Christianity	Europe	High	1.650	1.832	0.022	0.240	0.062	-0.120
Luxembourg	Christianity	Europe	High	2.310	2.099	0.054	0.357	-0.179	0.032
Macedonia	Christianity	Europe	Upper middle	1.873	1.938	0.007	0.121	0.005	-0.061
Malaysia	Islam	Asia	Upper middle	2.422	2.127	0.173	0.588	-0.294	0.001
Mali	Islam	SSA	Low	2.250	1.948	0.065	0.416	-0.094	0.208
Mexico	Christianity	Americas	Upper middle	2.342	1.925	0.090	0.460	-0.230	0.187
Moldova	Christianity	Europe	Lower middle	1.578	1.588	0.002	0.074	-0.014	-0.024
Montenegro	Christianity	Europe	Upper middle	1.869	1.916	0.023	0.244	-0.037	-0.085
Morocco	Islam	GME	Lower middle	2.028	1.845	0.021	0.236	-0.118	0.065
Netherlands	Christianity	Europe	High	2.335	1.943	0.093	0.496	-0.248	0.144
New Zealand	Christianity	Oceania	High	2.314	2.103	0.050	0.347	-0.174	0.037
Nigeria	Mixed	SSA	Lower middle	2.369	1.906	0.125	0.572	-0.177	0.286
Norway	Christianity	Europe	High	2.268	2.114	0.064	0.374	-0.170	-0.017
Pakistan	Islam	Asia	Lower middle	2.183	2.055	0.024	0.229	-0.014	0.115

Continued on next page

Table 4.8 Continued

Country	Religion	Area	Development	*Reported*	*Corrected*	Misclassification Error			
						Euclidean	*Rectilinear*	*Level 1*	*Level 3*
Philippines	Christianity	Asia	Lower middle	2.331	2.129	0.023	0.246	-0.123	0.080
Poland	Christianity	Europe	High	1.993	1.824	0.104	0.490	-0.207	-0.038
Portugal	Christianity	Europe	High	1.854	1.726	0.038	0.283	-0.135	-0.006
Puerto Rico	Christianity	Americas	High	2.444	2.083	0.072	0.430	-0.215	0.146
Qatar	Islam	GME	High	2.535	2.239	0.052	0.370	-0.112	0.185
Russia	Christianity	Europe	Upper middle	1.676	1.701	0.001	0.054	-0.001	-0.026
Rwanda	Christianity	SSA	Low	2.161	1.836	0.075	0.446	-0.223	0.102
Serbia	Christianity	Europe	Upper middle	1.876	1.738	0.086	0.451	-0.182	-0.044
Singapore	Confucian	Asia	High	2.289	1.944	0.082	0.469	-0.110	0.234
Slovakia	Christianity	Europe	High	1.706	1.920	0.052	0.352	0.038	-0.176
Slovenia	Christianity	Europe	High	1.893	1.785	0.009	0.156	-0.078	0.030
South Africa	Christianity	SSA	Upper middle	2.113	2.038	0.003	0.094	-0.028	0.047
Spain	Christianity	Europe	High	2.048	1.778	0.143	0.536	-0.268	0.001
Sweden	Christianity	Europe	High	2.333	2.146	0.033	0.287	-0.144	0.043
Switzerland	Christianity	Europe	High	2.334	2.273	0.011	0.159	0.010	0.070
Taiwan	Confucian	Asia	High	2.151	1.836	0.152	0.576	-0.288	0.027
Tanzania	Islam	SSA	Low	2.496	1.915	0.173	0.632	-0.266	0.316
Trinidad&Tobago	Mixed	Americas	High	2.399	2.099	0.070	0.430	-0.085	0.215
Tunisia	Islam	GME	Lower middle	1.972	1.868	0.013	0.176	-0.088	0.016

Continued on next page

Table 4.8 Continued

Country	Religion	Area	Development	*Reported*	*Corrected*	Misclassification Error			
						Euclidean	*Rectilinear*	*Level 1*	*Level 3*
Turkey	Islam	GME	Upper middle	2.145	1.909	0.145	0.559	0.022	0.258
Ukraine	Christianity	Europe	Lower middle	1.795	1.675	0.052	0.347	-0.147	-0.026
United States	Christianity	Americas	High	2.304	2.038	0.126	0.512	-0.256	0.010
Venezuela	Christianity	Americas	Upper middle	2.462	1.937	0.149	0.613	-0.219	0.306
Zimbabwe	Christianity	SSA	Low	2.000	2.088	0.005	0.112	0.032	-0.056

Notes: SSA, Sub-Saharan Africa. GME, Greater Middle East

Categorization of Development Stage is obtained from World Bank.

Major Religion and Area are from World Factbook.

Data Source: Identified results from section 4.3.

4.6.2 Proof of the Identification Procedure

In this section, we provide the formal proof that following Theorem 1 in Hu (2008), Assumptions 1 through 5 guarantee the nonparametric identification and estimation for the misclassification matrix $\mathbf{H}_{X|X^*,\mathbf{c}}$ through eigen-decomposition.

Proof. By the law of total probability and Assumptions 1 - 2, we have

$$\begin{aligned} Pr(Y,X,Z|\mathbf{c}) &= \sum_{X^*} Pr(Y,X,Z|X^*,\mathbf{c})Pr(X^*|\mathbf{c}) \\ &= \sum_{X^*} Pr(Y|X,Z,X^*,\mathbf{c})Pr(X|Z,X^*,\mathbf{c})Pr(Z|X^*,\mathbf{c})Pr(X^*|\mathbf{c}) \\ &= \sum_{X^*} Pr(Y|X^*,\mathbf{c})Pr(X|X^*,\mathbf{c})Pr(X^*,Z|\mathbf{c}) \end{aligned} \tag{4.16}$$

and

$$\begin{aligned} Pr(X,Z|\mathbf{c}) &= \sum_{X^*} Pr(X,Z|X^*,\mathbf{c})Pr(X^*|\mathbf{c}) = \sum_{X^*} Pr(X|X^*,Z,\mathbf{c})Pr(Z|X^*,\mathbf{c})Pr(X^*|\mathbf{c}) \\ &= \sum_{X^*} Pr(X|X^*,\mathbf{c})Pr(X^*,Z|\mathbf{c}) \end{aligned} \tag{4.17}$$

Defining in matrix, we have

$\mathbf{H}_{X|X^*,\mathbf{c}} \equiv [Pr(X=i|X^*=k,\mathbf{c})]_{i,k}$

$$= \left\{ \begin{array}{ccc} Pr(X=1|X^*=1,\mathbf{c}) & Pr(X=1|X^*=2,\mathbf{c}) & Pr(X=1|X^*=3,\mathbf{c}) \\ Pr(X=2|X^*=1,\mathbf{c}) & Pr(X=2|X^*=2,\mathbf{c}) & Pr(X=2|X^*=3,\mathbf{c}) \\ Pr(X=3|X^*=1,\mathbf{c}) & Pr(X=3|X^*=2,\mathbf{c}) & Pr(X=3|X^*=3,\mathbf{c}) \end{array} \right\} \tag{4.18}$$

$\mathbf{H}_{X^*,Z|\mathbf{c}} \equiv [Pr(X^*=i,Z=k|\mathbf{c})]_{i,k}$

$$= \left\{ \begin{array}{ccc} Pr(X^*=1,Z=1|\mathbf{c}) & Pr(X^*=1,Z=2|\mathbf{c}) & Pr(X^*=1,Z=3|\mathbf{c}) \\ Pr(X^*=2,Z=1|\mathbf{c}) & Pr(X^*=2,Z=2|\mathbf{c}) & Pr(X^*=2,Z=3|\mathbf{c}) \\ Pr(X^*=3,Z=1|\mathbf{c}) & Pr(X^*=3,Z=2|\mathbf{c}) & Pr(X^*=3,Z=3|\mathbf{c}) \end{array} \right\} \tag{4.19}$$

$\mathbf{H}_{X,Z|\mathbf{c}} \equiv [Pr(X=i, Z=k|\mathbf{c})]_{i,k}$

$$= \left\{ \begin{array}{ccc} Pr(X=1,Z=1|\mathbf{c}) & Pr(X=1,Z=2|\mathbf{c}) & Pr(X=1,Z=3|\mathbf{c}) \\ Pr(X=2,Z=1|\mathbf{c}) & Pr(X=2,Z=2|\mathbf{c}) & Pr(X=2,Z=3|\mathbf{c}) \\ Pr(X=3,Z=1|\mathbf{c}) & Pr(X=3,Z=2|\mathbf{c}) & Pr(X=3,Z=3|\mathbf{c}) \end{array} \right\} \tag{4.20}$$

$\mathbf{H}_{Y|Z,\mathbf{c}} \equiv [Pr(Y=i|Z=k, \mathbf{c})]_{i,k}$

$$= \left\{ \begin{array}{ccc} Pr(Y=1|Z=1,\mathbf{c}) & Pr(Y=1|Z=2,\mathbf{c}) & Pr(Y=1|Z=3,\mathbf{c}) \\ Pr(Y=2|Z=1,\mathbf{c}) & Pr(Y=2|Z=2,\mathbf{c}) & Pr(Y=2|Z=3,\mathbf{c}) \\ Pr(Y=3|Z=1,\mathbf{c}) & Pr(Y=3|Z=2,\mathbf{c}) & Pr(Y=3|Z=3,\mathbf{c}) \end{array} \right\} \tag{4.21}$$

$\mathbf{H}_{Y=3,X,Z|\mathbf{c}} \equiv [Pr(Y=3, X=i, Z=k|\mathbf{c})]_{i,k}$

$$= \left\{ \begin{array}{ccc} Pr(Y=3,X=1,Z=1|\mathbf{c}) & Pr(Y=3,X=1,Z=2,\mathbf{c}) & Pr(Y=3,X=1,Z=3|\mathbf{c}) \\ Pr(Y=3,X=2,Z=1|\mathbf{c}) & Pr(Y=3,X=2,Z=2,\mathbf{c}) & Pr(Y=3,X=2,Z=3|\mathbf{c}) \\ Pr(Y=3,X=3,Z=1|\mathbf{c}) & Pr(Y=3,X=3,Z=2,\mathbf{c}) & Pr(Y=3,X=3,Z=3|\mathbf{c}) \end{array} \right\} \tag{4.22}$$

The diagonal matrix is defined as $\mathbf{H}_{Y=3|X^*,\mathbf{c}} \equiv [Pr(Y=3|X^*=k, \mathbf{c})]_k$

$$= \left\{ \begin{array}{ccc} Pr(Y=3|X^*=1,\mathbf{c}) & 0 & 0 \\ 0 & Pr(Y=3|X^*=2,\mathbf{c}) & 0 \\ 0 & 0 & Pr(Y=3|X^*=3,\mathbf{c}) \end{array} \right\} \tag{4.23}$$

Expressing the relationship in Equations 4.16 and 4.17 with matrix definition, we have

$$\mathbf{H}_{Y=3,X,Z|\mathbf{c}} = \mathbf{H}_{X|X^*,\mathbf{c}} \mathbf{H}_{Y=3|X^*,\mathbf{c}} \mathbf{H}_{X^*,Z|\mathbf{c}} \tag{4.24}$$

and

$$\mathbf{H}_{X,Z|\mathbf{c}} = \mathbf{H}_{X|X^*,\mathbf{c}} \mathbf{H}_{X^*,Z|\mathbf{c}} \tag{4.25}$$

Given Assumption 3, $\mathbf{H}_{X,Z|\mathbf{c}}$ is invertible. Post-multiplying both sides of Equation 4.24 by $\mathbf{H}_{X,Z|\mathbf{c}}^{-1}$ yields the following identification equation:

$$\begin{aligned} \mathbf{H}_{Y=3,X,Z|\mathbf{c}}\mathbf{H}_{X,Z|\mathbf{c}}{}^{-1} &= \mathbf{H}_{X|X^*,\mathbf{c}}\mathbf{H}_{Y=3|X^*,\mathbf{c}}\mathbf{H}_{X^*,Z|\mathbf{c}}\mathbf{H}_{X^*,Z|\mathbf{c}}{}^{-1}\mathbf{H}_{X|X^*,\mathbf{c}}{}^{-1} \\ &= \mathbf{H}_{X|X^*,\mathbf{c}}\mathbf{H}_{Y=3|X^*,\mathbf{c}}\mathbf{H}_{X|X^*,\mathbf{c}}{}^{-1} \end{aligned} \tag{4.26}$$

Consequently,

$$\mathbf{H}_{X|X^*,\mathbf{c}} = \zeta(\mathbf{H}_{Y=3,X,Z|\mathbf{c}}\mathbf{H}_{X,Z|\mathbf{c}}{}^{-1}) \tag{4.27}$$

where $\zeta(\cdot)$ denotes the mapping from a square matrix to its eigenvector matrix. The three eigenvalues are the three elements of the diagonal matrix $\mathbf{H}_{Y=3|X^*,\mathbf{c}}$. Note that in our case each eigenvector is a distribution, the summation of the elements is equal to 1, which indicates that the $\zeta(\cdot)$ is normalized. The monotonically ordered eigenvalues in X^*, implied in Assumption 4 and 5, guarantee a unique mapping of $\zeta(\cdot)$. Hence, $\mathbf{H}_{Y=3,X,Z|\mathbf{c}}\mathbf{H}_{X,Z|\mathbf{c}}{}^{-1}$ is obtained by direct calculation. Following Hu (2008), $\mathbf{H}_{X|X^*,\mathbf{c}}$ can be uniquely identified.

Pre-multiplying the observed probability distribution of reported SWB by the inverse of $\mathbf{H}_{X|X^*,\mathbf{c}}$ yields $[Pr(X^* = k|\mathbf{c})]_k = \mathbf{H}_{X|X^*,\mathbf{c}}{}^{-1}\,[Pr(X = k|\mathbf{c})]_k$, which is the true latent status of SWB.

$$\begin{aligned} &= \left\{ \begin{array}{ccc} Pr(X=1|X^*=1,\mathbf{c}) & \cdots & Pr(X=1|X^*=3,\mathbf{c}) \\ Pr(X=2|X^*=1,\mathbf{c}) & \cdots & Pr(X=2|X^*=3,\mathbf{c}) \\ Pr(X=3|X^*=1,\mathbf{c}) & \cdots & Pr(X=3|X^*=3,\mathbf{c}) \end{array} \right\}^{-1} \left\{ \begin{array}{c} Pr(X=1|\mathbf{c}) \\ Pr(X=2|\mathbf{c}) \\ Pr(X=3|\mathbf{c}) \end{array} \right\} \\ &= \left\{ \begin{array}{c} Pr(X^*=1|\mathbf{c}) \\ Pr(X^*=2|\mathbf{c}) \\ Pr(X^*=3|\mathbf{c}) \end{array} \right\} \end{aligned} \tag{4.28}$$

$\square$

QED.

4.6.3 Evaluation of Assumptions 3 and 4

Assumption 3 requires matrix $\mathbf{H}_{X,Z|\mathbf{c}}$ to satisfy the full rank condition. We rigorously test this in a sequential manner following the procedure proposed by Robin and Smith (2000). Specifically, we first test the null hypothesis of $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) = 0$ against the alternative hypothesis of $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) > 0$. If the null is rejected at 5% significant level, we then test the null of $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) = 1$ against the alternative of $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) > 1$. If the null is again rejected at 5% level, we then test the null of $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) = 2$ against the alternative of $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) > 2$. As we have three levels in each measurement of SWB, the dimension of matrix $\mathbf{H}_{X,Z|\mathbf{c}}$ is accordingly 3×3. If the null is still rejected at 5% significant level, then the full rank condition is satisfied.

By the full factorials of two gender, three age and three income categories, we divide the data of each country into 18 sub-samples. We apply the rank test for each of the 18 demographic groups across all the 80 countries with 1000 bootstrapping replications. Across the 1440 (80×18) tests implemented, the null hypotheses of $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) = 0$ and $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) = 1$ are rejected at 1% significant level. Due to the limitations of sample size, the null hypothesis of $Rank(\mathbf{H}_{X,Z|\mathbf{c}}) = 2$ can not be rejected at 5% significant level for some demographic groups in some countries. In Assumption 4, we require the diagonal matrix $\mathbf{H}_{Y=3|X^*,\mathbf{c}}$ to have three distinct elements, demographic groups not satisfying this condition are also not identified. Collectively, 107 out of 1440 demographic groups are not identifiable and removed from the data set. Two-sided Binomial tests for the hypothesis that each identified group is chosen with equal probability (1/18) can not be rejected for all 18 demographic groups at 10% significant level (Table 4.9). Therefore, the proportion of each identified group observed after the sample attrition is not significantly different from random (with probability of 1/18).

Table 4.9: Results of Binomial Test for Each of the 18 Demographic Groups

	(1) Subgroup	(2) Subgroup	(3) Subgroup	(4) Subgroup	(5) Subgroup	(6) Subgroup	(7) Subgroup	(8) Subgroup	(9) Subgroup	(10) Subgroup	(11) Subgroup	(12) Subgroup	(13) Subgroup	(14) Subgroup	(15) Subgroup	(16) Subgroup	(17) Subgroup	(18) Subgroup
P-value	0.605	0.605	0.527	0.774	0.455	0.455	0.527	0.774	0.527	0.605	0.605	0.185	0.687	0.863	0.388	0.527	0.774	0.120
Observed	75	75	74	77	73	73	74	77	74	75	75	68	76	78	72	74	77	66

Notes: N = 1,440. The null hypothesis is that the probability of observing each identified demographic group is equal to 1/18

4.6.4 Misclassification Errors for Five Representative Countries: By Demographic Groups

Table 4.10 to 4.12 present the misclassification errors aggregated at different demographic groups (income, gender and age) for five representative countries, including China, Germany, Iraq, Mali, and the United States.

Table 4.10: Misclassification Errors, By Income Group *(five representative countries)*

Country	**Level 1** (Unhappy)			**Level 3** (Very Happy)		
	reported	*corrected*	*difference*	*reported*	*corrected*	*difference*
Panel A: Low Income						
China(Mainland)	0.126	0.295	-0.170	0.205	0.311	-0.106
	(0.013)	(0.100)		(0.019)	(0.071)	
Germany	0.098	0.387	-0.289	0.192	0.268	-0.076
	(0.009)	(0.101)		(0.011)	(0.084)	
Iraq	0.113	0.214	-0.101	0.292	0.274	0.018
	(0.014)	(0.101)		(0.030)	(0.090)	
Mali	0.073	0.099	-0.026	0.623	0.353	0.269
	(0.020)	(0.106)		(0.037)	(0.108)	
United States	0.036	0.201	-0.165	0.445	0.521	-0.076
	(0.005)	(0.118)		(0.013)	(0.080)	
Panel B: Middle Income						
China(Mainland)	0.301	0.330	-0.029	0.190	0.301	-0.110
	(0.010)	(0.108)		(0.008)	(0.071)	
Germany	0.231	0.392	-0.161	0.142	0.227	-0.085
	(0.009)	(0.089)		(0.008)	(0.047)	
Iraq	0.490	0.369	0.122	0.074	0.155	-0.082
	(0.012)	(0.091)		(0.006)	(0.054)	
Mali	0.348	0.201	0.146	0.266	0.231	0.035
	(0.031)	(0.105)		(0.029)	(0.104)	
United States	0.140	0.464	-0.324	0.307	0.324	-0.017
	(0.009)	(0.096)		(0.012)	(0.055)	
Panel C: High Income						
China(Mainland)	0.165	0.208	-0.043	0.197	0.182	0.015
	(0.006)	(0.103)		(0.007)	(0.047)	
Germany	0.127	0.356	-0.229	0.199	0.308	-0.109
	(0.006)	(0.085)		(0.006)	(0.034)	
Iraq	0.357	0.511	-0.154	0.093	0.136	-0.044
	(0.009)	(0.095)		(0.005)	(0.051)	
Mali	0.116	0.318	-0.202	0.419	0.160	0.259
	(0.013)	(0.104)		(0.021)	(0.089)	
United States	0.067	0.326	-0.259	0.390	0.343	0.047
	(0.004)	(0.092)		(0.007)	(0.053)	

Notes: Numbers reported in parentheses are bootstrap standard errors based on 300 repetitions.
Source: Identified results from section 2.3.

Table 4.11: Misclassification Errors, By Gender Group *(five representative countries)*

Country	**Level 1** (Unhappy)			**Level 3** (Very Happy)		
	reported	*corrected*	*difference*	*reported*	*corrected*	*difference*
Panel A: Female						
China(Mainland)	0.190	0.249	-0.059	0.205	0.212	-0.008
	(0.007)	(0.111)		(0.007)	(0.056)	
Germany	0.164	0.300	-0.137	0.189	0.288	-0.099
	(0.006)	(0.071)		(0.006)	(0.032)	
Iraq	0.341	0.318	0.023	0.107	0.102	0.006
	(0.009)	(0.089)		(0.006)	(0.045)	
Mali	0.149	0.258	-0.109	0.408	0.120	0.288
	(0.015)	(0.100)		(0.022)	(0.078)	
United States	0.075	0.402	-0.327	0.390	0.345	0.045
	(0.004)	(0.082)		(0.008)	(0.059)	
Panel B: Male						
China(Mainland)	0.228	0.269	-0.041	0.188	0.258	-0.070
	(0.007)	(0.087)		(0.007)	(0.048)	
Germany	0.140	0.451	-0.311	0.173	0.269	-0.096
	(0.006)	(0.091)		(0.006)	(0.043)	
Iraq	0.413	0.580	-0.166	0.094	0.199	-0.105
	(0.010)	(0.100)		(0.006)	(0.059)	
Mali	0.178	0.257	-0.079	0.420	0.291	0.129
	(0.017)	(0.099)		(0.022)	(0.096)	
United States	0.080	0.266	-0.186	0.373	0.397	-0.024
	(0.005)	(0.095)		(0.008)	(0.044)	

Notes: Numbers reported in parentheses are bootstrap standard errors based on 300 repetitions.
Source: Identified results from section 2.3.

Table 4.12: Misclassification Errors, By Age Group *(five representative countries)*

Country	**Level 1** (Unhappy)			**Level 3** (Very Happy)		
	reported	*corrected*	*difference*	*reported*	*corrected*	*difference*
Panel A: Age Band 15-29 Years						
China(Mainland)	0.210	0.354	-0.144	0.193	0.283	-0.090
	(0.011)	(0.092)		(0.011)	(0.082)	
Germany	0.105	0.428	-0.323	0.198	0.259	-0.061
	(0.008)	(0.081)		(0.010)	(0.042)	
Iraq	0.339	0.402	-0.062	0.128	0.126	0.003
	(0.011)	(0.083)		(0.008)	(0.042)	
Mali	0.119	0.424	-0.305	0.456	0.196	0.260
	(0.016)	(0.118)		(0.025)	(0.090)	
United States	0.076	0.086	-0.010	0.371	0.330	0.041
	(0.007)	(0.114)		(0.012)	(0.073)	
Panel B: Age Band 30-49 Years						
China(Mainland)	0.205	0.110	0.095	0.202	0.225	-0.023
	(0.007)	(0.118)		(0.007)	(0.052)	
Germany	0.137	0.269	-0.133	0.195	0.316	-0.121
	(0.007)	(0.101)		(0.008)	(0.042)	
Iraq	0.388	0.481	-0.093	0.088	0.097	-0.009
	(0.010)	(0.119)		(0.006)	(0.047)	
Mali	0.169	0.119	0.049	0.398	0.229	0.170
	(0.017)	(0.114)		(0.024)	(0.114)	
United States	0.077	0.519	-0.442	0.380	0.373	0.007
	(0.005)	(0.117)		(0.009)	(0.053)	
Panel C: Age Band 50 and More Years						
China(Mainland)	0.221	0.462	-0.241	0.186	0.212	-0.027
	(0.009)	(0.117)		(0.009)	(0.061)	
Germany	0.190	0.437	-0.247	0.162	0.255	-0.093
	(0.008)	(0.092)		(0.007)	(0.046)	
Iraq	0.428	0.469	-0.041	0.079	0.340	-0.261
	(0.016)	(0.127)		(0.009)	(0.145)	
Mali	0.233	0.251	-0.018	0.374	0.177	0.197
	(0.030)	(0.132)		(0.034)	(0.090)	
United States	0.079	0.288	-0.209	0.390	0.393	-0.003
	(0.005)	(0.090)		(0.009)	(0.067)	

Notes: Numbers reported in parentheses are bootstrap standard errors based on 300 repetitions.
Source: Identified results from section 2.3.

5. CONCLUSIONS AND DISCUSSION

Public policy is essential to society. Good policy initiatives help society to operate more effectively without necessarily sacrificing equity. Quantitatively evaluating the efficiency and impact of public policies determines whether the policy making outcomes result as planned. While the intended effect is well-anticipated, the unintended effects are generally difficult to foresee. Both impacts are of public-interest and they are key for social welfare.

This dissertation covers three slightly different angles to study public policy. In the first essay, we identify potential unintended effects leading to an increase in the occurrence of violent crime rates following the initial passage of Alabama's anti-illegal immigrant bill—HB56. One explaining mechanism is provided in the seminal work of Gary Becker (Becker, 1968). HB56's strict restriction for undocumented immigrants taking job positions decreased the opportunity cost of criminal behavior. Another possible mechanism is associated to the growing literature on expressive value of law, which suggests the act of enacting a particular law serves the function of shaping norms or prescribed attitude towards behavior (Bursztyn and Jensen, 2017). The adoption of an anti-immigration law will send out a signal of increased tolerance to discrimination behavior against undocumented workers, which can also potentially generate tensions or conflicts that may result in violent crimes against undocumented immigrants. Whether the increase in violent crimes is committed by undocumented immigrants or against them, and through which underlying mechanism merit further investigation using richer data.

In the second essay, we examine the impact of an immigration law—AC 21 on high-skilled labor markets. The unexpected exemption of the H-1B visa cap for working in eligible academic institutions or research entities changed the risk of being denied entrance to the US labor market for foreign-born Ph.D. graduates. The finding suggests that AC 21 significantly boosted the preference for working in academia by 5% among US-trained foreign doctorate recipients. AC 21 also indirectly caused foreign Ph.D. graduates to be 3-4% less likely to begin a career in industry. These results further pass a series of variation in specifications and placebo and falsification tests. The potential implication of these findings is profound. Well-educated foreign professionals are playing leading roles in innovation and technology development in the United States. Being able to retain the human capital actively contributing to fundamental research can significantly benefit United States in the long-term. The results of this study indicate that this goal can be effectively achieved by simply reducing the barriers in the visa application process for high-skilled professionals.

When policy makers design intervention programs, their goals are usually measured by objectively measurable outcomes, such as food access and expenditures, changes in prices, income, health care access, etc. While objective measures serve the purpose of evaluating the efficacy of public policies, subjective well-being provides insights on whether a policy benefits the general public in a broader sense. Indeed, recent studies attempt to use subjective welfare to study poverty (Blank, 2008), health and the environment (Zhang et al., 2015), and social progress (Fehder et al., 2018). However, the inherent measurement error in reported SWB restricts it from being applied for public policy research.

In the third essay, we use a newly developed method to recover the latent true distribution of subjective well-being for 80 countries. The results show that SWB, collected in surveys, has substantial misclassification errors. Religious beliefs and the development stage of a country play critical roles in determining the magnitude of misclassification errors. Reexamining the Easterlin paradox and modified-Easterlin hypothesis implies that although *reported* SWB is not associated with GDP *per capita*, *corrected* measure of SWB is; and this relationship is robust beyond an income satiation point. To apply SWB for program evaluation study, future research still need to identify the influencing factors which causally impact the misclassification errors in reported SWB. A richer longitudinal dataset and exogenous variation would be ideal ways to move in this direction.

REFERENCES

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association 105*(490), 493–505.

Abadie, A., A. Diamond, and J. Hainmueller (2015). Comparative politics and the synthetic control method. *American Journal of Political Science 59*(2), 495–510.

Abadie, A. and J. Gardeazabal (2003). The economic costs of conflict: A case study of the basque country. *American economic review 93*(1), 113–132.

Addy, S. (2012). A cost-benefit analysis of the new alabama immigration law. *Center for Business and Economic Research, Culverhouse College of Commerce and Business Administration. University of Alabama.*.

Alesina, A. and P. Giuliano (2010). The power of the family. *Journal of Economic growth 15*(2), 93–125.

Amuedo-Dorantes, C. and D. Furtado (2017). Settling for Academia? H-1B Visas and the Career Choices of International Students in the United States. *Journal of Human Resources*, 0816–8167r1.

An, Y., M. R. Baye, Y. Hu, J. Morgan, and M. Shum (2017). Identification and estimation of online price competition with an unknown number of firms. *Journal of Applied Econometrics 32*(1), 80–102.

An, Y., Y. Hu, and M. Shum (2010). Estimating first-price auctions with an unknown number of bidders: A misclassification approach. *Journal of Econometrics 157*(2), 328–341.

Anderson, S. and M. D. Platzer (2006). *American Made: the impact of immigrant entrepreneurs and professionals on US competitiveness*. National Venture Capital Association.

Argyle, M. (2003). 18 causes and correlates of happiness. *Well-being: The foundations of hedonic psychology 353*.

Bailey, W. C. (1984). Poverty, inequality, and city homicide rates. *Criminology 22*(4), 531–550.

Baker, S. R. (2015). Effects of immigrant legalization on crime. *American Economic Review 105*(5), 210–13.

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The economic dimensions of crime*, pp. 13–68. Springer.

Bell, B., F. Fasani, and S. Machin (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and statistics 21*(3), 1278–1290.

Benjamin, D. J., O. Heffetz, M. S. Kimball, and A. Rees-Jones (2012, May). What do you think would make you happier? what do you think you would choose? *American Economic Review 102*(5), 2083–2110.

Blanchflower, D. G. and A. J. Oswald (2004). Money, sex and happiness: An empirical study. *Scandinavian Journal of Economics 106*(3), 393–415.

Blank, R. M. (2008). How to improve poverty measurement in the united states. *Journal of Policy Analysis and Management 27*(2), 233–254.

Blau, J. R. and P. M. Blau (1982). The cost of inequality: Metropolitan structure and violent crime. *American sociological review*, 114–129.

Bohn, S., M. Lofstrom, and S. Raphael (2014). Did the 2007 legal arizona workers act reduce the state's unauthorized immigrant population? *Review of Economics and Statistics 96*(2), 258–269.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics 118*(4), 1335–1374.

Borjas, G. J. (2005, May). The labor-market impact of high-skill immigration. *American Economic Review 95*(2), 56–60.

Borjas, G. J. and K. B. Doran (2012). The collapse of the Soviet Union and the productivity of American mathematicians. *Quarterly Journal of Economics 127*(3), 1143–1203.

Bursztyn, L. and R. Jensen (2017). Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure. *Annual Review of Economics 9*, 131–153.

Butcher, K. F. and A. M. Piehl (1998). Cross-city evidence on the relationship between immigration and crime. *Journal of policy analysis and management*, 457–493.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Campante, F. and D. Yanagizawa-Drott (2015). Does religion affect economic growth and happiness? evidence from ramadan. *The Quarterly Journal of Economics 130*(2), 615–658.

Chalfin, A. (2013). What is the contribution of mexican immigration to us crime rates? evidence from rainfall shocks in mexico. *American Law and Economics Review 16*(1), 220–268.

Cheng, C. and M. Hoekstra (2013). Does strengthening self-defense law deter crime or escalate violence? Evidence from expansions to castle doctrine. *Journal of Human Resources 48*(3), 821–854.

Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. *The American Economic Review 105*(5), 1–33.

Clark, A. E., P. Frijters, and M. A. Shields (2008, March). Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles. *Journal of Economic Literature 46*(1), 95–144.

Cohen, W. M., R. R. Nelson, and J. P. Walsh (2002). Links and impacts: the influence of public research on industrial R&D. *Management science 48*(1), 1–23.

Deaton, A. (2008). Income, health, and well-being around the world: Evidence from the gallup world poll. *Journal of Economic Perspectives 22*(2), 53–72.

Deaton, A. and A. A. Stone (2013). Two happiness puzzles. *The American economic review 103*(3), 591–597.

Di Tella, R. and R. MacCulloch (2008, December). Happiness adaptation to income beyond "basic needs". Working Paper 14539, National Bureau of Economic Research.

Di Tella, R., J. H.-D. New, and R. MacCulloch (2010). Happiness adaptation to income and to status in an individual panel. *Journal of Economic Behavior & Organization 76*(3), 834 – 852.

Di Tella, R. and E. Schargrodsky (2004). Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review 94*(1), 115–133.

Diener, E. (2000). Subjective well-being. the science of happiness and a proposal for a national index. *American Psychologist 55*(1), 34–43.

Diener, E., R. Inglehart, and L. Tay (2013). Theory and validity of life satisfaction scales. *Social Indicators Research 112*(3), 497–527.

Dolan, P., T. Peasgood, and M. White (2008). Do we really know what makes us happy? a review of the economic literature on the factors associated with subjective well-being. *Journal of economic psychology 29*(1), 94–122.

Easterlin, R. A. (1974). Does economic growth improve the human lot? some empirical evidence. *Nations and households in economic growth 89*, 89–125.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization 27*(1), 35 – 47.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of political Economy 81*(3), 521–565.

Ellison, C. (1994). Religion, the life stress paradigm, and the study of depression. *Religion in Aging and Health*, 78.

Fajnzylber, P., D. Lederman, and N. Loayza (2002). What causes violent crime? *European economic review 46*(7), 1323–1357.

Fausset, R. (2011). Alabama enacts anti-illegal-immigration law described as nationâĂŹs strictest. *Los Angeles Times*.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015, October). The next generation of the penn world table. *American Economic Review 105*(10), 3150–82.

Fehder, D., M. Porter, and S. Stern (2018). The empirics of social progress: The interplay between subjective well-being and societal performance. In *AEA Papers and Proceedings*, Volume 108, pp. 477–82.

Feng, S. and Y. Hu (2013). Misclassification errors and the underestimation of the US unemployment rate. *The American Economic Review 103*(2), 1054–1070.

Fleisher, B. M. (1966). The effect of income on delinquency. *The American Economic Review 56*(1/2), 118–137.

Fleurbaey, M. (2009, December). Beyond gdp: The quest for a measure of social welfare. *Journal of Economic Literature 47*(4), 1029–75.

Francis, L. (2010). Religion and happiness: perspectives from the psychology of religion, positive psychology and empirical theology.

Frey, B. S. and A. Stutzer (2000). Happiness, economy and institutions. *The Economic Journal 110*(466), 918–938.

Frey, B. S. and A. Stutzer (2002, June). What can economists learn from happiness research? *Journal of Economic Literature 40*(2), 402–435.

Frey, B. S. and A. Stutzer (2010). *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*. Princeton University Press.

Galor, O. and Ö. Özaka (2016). The agricultural origins of time preference. *The American Economic Review 106*(10), 3064–3103.

Hagedorn, J. W. (1996). Happiness and self-deception: An old question examined by a new measure of subjective well-being. *Social Indicators Research 38*(2), 139–160.

Hagerty, M. R. and R. Veenhoven (2003). Wealth and happiness revisited – growing national income does go with greater happiness. *Social Indicators Research 64*(1), 1–27.

Helliwell, J., R. Layard, and J. Sachs (2012). World happiness report.

Hoekstra, M. and S. Orozco-Aleman (2017). Illegal immigration, state law, and deterrence. *American Economic Journal: Economic Policy 9*(2), 228–52.

Horan, P. M. and P. L. Austin (1974). The social bases of welfare stigma. *Social problems 21*(5), 648–657.

Hu, Y. (2008). Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution. *Journal of Econometrics 144*(1), 27–61.

Hunt, J. (2011). Which immigrants are most innovative and entrepreneurial? Distinctions by entry visa. *Journal of Labor Economics 29*(3), 417–457.

Hunt, J. (2015). Are Immigrants the Most Skilled US Computer and Engineering Workers? *Journal of Labor Economics 33*(S1), S39–S77.

Hunt, J. and M. Gauthier-Loiselle (2010). How much does immigration boost innovation? *American Economic Journal: Macroeconomics 2*(2), 31–56.

Islam, A. (2009). The substitutability of labor between immigrants and natives in the Canadian labor market: circa 1995. *Journal of Population Economics 22*(1), 199–217.

Islam, A., F. Islam, and C. Nguyen (2017). Skilled Immigration, Innovation and Wages of Native-born American. *Industrial Relations: A Journal of Economy and Society 56*(3), 459–488.

John, P. and K. F. Zimmermann (1994). Native wage impacts of foreign labor: a random effects panel analysis. *Journal of Population Economics 7*(2), 177–192.

Kahneman, D. and A. B. Krueger (2006, March). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives 20*(1), 3–24.

Kato, T. and C. Sparber (2013). Quotas and quality: The effect of H-1B visa restrictions on the pool of prospective undergraduate students from abroad. *Review of Economics and Statistics 95*(1), 109–126.

Kerr, S. P., W. R. Kerr, and W. F. Lincoln (2015). Skilled immigration and the employment structures of US firms. *Journal of Labor Economics 33*(S1), S147–S186.

Kerr, W. R. (2013). U.S. high-skilled immigration, innovation, and entrepreneurship: empirical approaches and evidence. *NBER Working Paper No. 19377*.

Kerr, W. R. and W. F. Lincoln (2010). The supply side of innovation: H-1B visa reforms and us ethnic invention. *Journal of Labor Economics 28*(3), 473–508.

Lan, X. (2012). Permanent visas and temporary jobs: evidence from postdoctoral participation of foreign Ph.Ds in the united states. *Journal of Policy Analysis and Management 31*(3), 623–640.

Lan, X. (2013). The effects of green cards on the wages and innovations of new Ph.Ds. *Journal of Policy Analysis and Management 32*(4), 807–834.

Levitt, S. D. (2002). Using electoral cycles in police hiring to estimate the effects of police on crime: Reply. *American Economic Review 92*(4), 1244–1250.

Mastrobuoni, G. and P. Pinotti (2015). Legal status and the criminal activity of immigrants. *American Economic Journal: Applied Economics 7*(2), 175–206.

Mithas, S. and H. C. Lucas Jr (2010). Are foreign it workers cheaper? U.S. visa policies and compensation of information technology professionals. *Management Science 56*(5), 745–765.

Moffitt, R. (1983). An economic model of welfare stigma. *The American Economic Review 73*(5), 1023–1035.

Moser, P., A. Voena, and F. Waldinger (2014). German-Jewish emigrés and U.S. invention. *American Economic Review 104*(10), 3222–55.

OECD (2013). OECD Guidelines on Measuring Subjective Well-being.

Ottaviano, G. I. and G. Peri (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association 10*(1), 152–197.

Peri, G. (2007). *Higher education, innovation and growth*. Education and Training in Europe, Oxford: Oxford University Press.

Peri, G., K. Shih, and C. Sparber (2014a). The effects of foreign skilled workers on natives: Evidence from the H-1B visa lottery. *UC Davis Working Paper*.

Peri, G., K. Shih, and C. Sparber (2015). STEM workers, H-1B visas, and productivity in US cities. *Journal of Labor Economics 33*(S1), S225–S255.

Peri, G., K. Y. Shih, and C. Sparber (2014b). Foreign STEM workers and native wages and employment in U.S. cities. Technical report, National Bureau of Economic Research.

Pinotti, P. (2017). Clicking on heaven's door: The effect of immigrant legalization on crime. *American Economic Review 107*(1), 138–68.

Robin, J.-M. and R. J. Smith (2000). Tests of rank. *Econometric Theory 16*(02), 151–175.

Sacks, D. W., B. Stevenson, and J. Wolfers (2010). Subjective well-being, income, economic development and growth. Working Paper 16441, National Bureau of Economic Research.

Sheridan, Z., P. Boman, A. Mergler, and M. J. Furlong (2015). Examining well-being, anxiety, and self-deception in university students. *Cogent Psychology 2*(1), 993850.

Shih, K. (2016). Labor Market Openness, H-1B Visa Policy, and the Scale of International Student

Enrollment in the United States. *Economic Inquiry 54*(1), 121–138.

Spenkuch, J. L. (2013). Understanding the impact of immigration on crime. *American law and economics review 16*(1), 177–219.

Stephan, P. and J. Ma (2005). The increased frequency and duration of the postdoctorate career stage. *American Economic Review*, 71–75.

Stevenson, B. and J. Wolfers (2008a, August). Economic growth and subjective well-being: Reassessing the easterlin paradox. Working paper, National Bureau of Economic Research.

Stevenson, B. and J. Wolfers (2008b). Happiness inequality in the united states. *The Journal of Legal Studies 37*(S2), S33–S79.

Stevenson, B. and J. Wolfers (2013, May). Subjective well-being and income: Is there any evidence of satiation? *American Economic Review 103*(3), 598–604.

Stuen, E. T., A. M. Mobarak, and K. E. Maskus (2012). Skilled immigration and innovation: evidence from enrollment fluctuations in US doctoral programs. *The Economic Journal 122*(565), 1143–1176.

Thurgood, L. (2004). Graduate enrollment in science and engineering fields reaches new peak; first-time enrollment of foreign students declines. *InfoBrief NSF 4*, 326.

U.S. Congress, U. (2000). American Competitiveness in the Twenty-first Century Act of 2000.

Wadhwa, V., A. Saxenian, B. A. Rissing, and G. Gereffi (2007). America's new immigrant entrepreneurs: Part i.

Zhang, X., X. Zhang, and X. Chen (2015). Happiness in the air: How does dirty sky affect subjective well-being?