May 2019

Three Essays on Public Policy: Physician Labor Supply, Food Insecurity, and Income Inequality

Judith Liu
Syracuse University

Abstract

This dissertation comprises three papers on physician labor supply, food insecurity, and income inequality. My research broadly explores how public policies and government programs affect individual behavior and how effectively they alleviate inequality and poverty.

Chapter 1 estimates the impact of a transitory reduction in hours during physicians' early career on their long-term labor supply. I exploit the work-hour regulations that limit the maximum workweek by residents as the source of exogenous variation. The results show that exposure to the regulations significantly decreases practicing physicians' labor supply by about four hours per week on average, with female physicians being more responsive to a given reduction in early career hours. Distributional results using a changes-in-changes model confirm that the regulations primarily affect the upper end of the work hours distribution. To reveal potential mechanisms of these effects, I find that the reform increases the probabilities of marriage and having a child, as well as the total number of children, for female physicians. In contrast, it does not have a significant impact on marriage and fertility outcomes for male physicians. These findings provide a better understanding of physicians' hours of work in response to the reform over time and the role of gender with respect to labor supply behavior and family formation decisions.

Chapter 2 studies the role of government programs in alleviating differential exposure to food insecurity. We provide a framework that conceptualizes how the Supplemental Nutrition Assistance Program (SNAP) could have differences in benefit levels across racial/ethnic groups. We decompose differences in SNAP benefit levels into three components: differences in eligibility, participation, and generosity. We then link the results to differences in food consumption to provide implications on food insecurity differentials. Our results reveal that

SNAP has different pathways to reducing food insecurity for different populations. Among the three components, eligibility contributes the most to SNAP benefits for both blacks and Hispanics relative to whites. However, SNAP reduces differences in food consumption between blacks/Hispanics and whites by a modest amount, which is likely not enough to reduce the differences in the resource gaps between groups. We also provide an exploratory analysis of how changes to SNAP policy rules might affect differences in food insecurity across groups. Our results suggest that the automatic enrollment policy might be effective in ameliorating the disparities.

Chapter 3 estimates the effects of trade liberalization on household income inequality and investigates whether trade liberalization or domestic reforms are the main influence factors of the rising inequality since 1980 in Taiwan, a middle-income open economy. We construct an empirical model by decomposing the sources of household disposable income in the quintile ratio. Using time-series data from 1980 to 2015 to estimate the long-run effect, we find that trade liberalization raises income inequality overall. When separating trade partners into OECD and non-OECD countries, our results show that net exports to OECD countries increase inequality, whereas net exports to non-OECD countries insignificantly decrease inequality. Moreover, we provide evidence that domestic reforms, particularly technological progress in favor of skilled labor and industrial structural change, rather than trade liberalization, are the main driving forces of income inequality.

THREE ESSAYS ON PUBLIC POLICY: PHYSICIAN LABOR SUPPLY, FOOD INSECURITY, AND INCOME INEQUALITY

by

Judith Liu

B.A., National Chengchi University, 2010
M.A., National Taiwan University, 2012
M.A., Washington University in St. Louis, 2014

Dissertation
Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in *Economics*.

Syracuse University
May 2019

Acknowledgments

I would like to express my sincerest gratitude to my advisor Alfonso Flores-Lagunes for his invaluable guidance and support. His breadth of knowledge, competent and patient guidance, and thoughtful criticism have made my graduate study fruitful and productive. I am also thankful to my dissertation defense committee: Gary Engelhardt, Hugo Jales, Jeffrey Kubik, Bill Horrace, and Colleen Helfin, for their helpful advice and feedback. A further thank you goes to Elizabeth Ashby, Jerry Evensky, Carmen Carrión-Flores, Michael Wasylenko, the faculty and fellow graduate students at the Department of Economics, Aging Studies Institute, and the Center for Policy Research. Lastly, I dedicate this dissertation to my parents and my brother who have always supported me unconditionally in my endeavors.

Table of Contents

1. The Long-Term Impact of Work-Hour Regulations on Physician Labor Supply..............1

1.1 Introduction...2

1.2 Physician Work-Hour Regulations in the United States 6

1.3 Conceptual Framework 8

1.4 Data and the Effectiveness of the Reform 12

1.4.1 Data Construction and Summary Statistics...12

1.4.2 First Stage: The Effectiveness of the Reform...15

1.5 The Impact on Long-Term Labor Supply 17

1.5.1 Difference-in-Differences Approach 17

1.5.2 Estimated Mean Effects 20

1.5.3 Potential Threats to Identification...22

1.5.4 Distributional Effects 25

1.6 Mechanisms 27

1.7 Conclusions...30

References...33

Appendix...46

2. Differential Exposure to Food Insecurity and Government Policy 52

2.1 Introduction...53

2.2 Background...56

2.3 Decomposition of SNAP 59

2.3.1 Baseline model..59

2.3.2 Numerical Example 61

2.3.3 General Setting...63

2.3.4 Linkage to Differential Exposure to Food Insecurity 66

2.4 Data 68

2.5 Results..71

2.5.1 Descriptive Statistics 71

2.5.2 Decomposition Results 72

2.6 Conclusions....74
References....77

3. The Effects of Trade Liberalization and Domestic Reforms on Income Inequality in a Middle-Income Open Economy: Evidence from Taiwan 85
3.1 Introduction....86
3.2 Literature Review....90
3.3 Income Inequality, Trade Liberalization, and Domestic Reforms 93
3.3.1 Household Income Inequality....93
3.3.2 Trade Liberalization....95
3.3.3 Industrial Development Policy: Structural Change and Technological Progress....96
3.3.4 Reforms in Higher Education....99
3.3.5 Financial Market Reforms....100
3.3.6 Changes in Social Welfare and Social Insurance policies....101
3.4 Empirical Model....102
3.4.1 Influence Factors of R_W....103
3.4.2 Influence Factors of R_A....106
3.4.3 Influence Factors of R_{GT}....106
3.4.4 Empirical Model and Expected Effects....107
3.5 Empirical Results....110
3.5.1 Data, Unit Root Test, and Methodology....110
3.5.2 Estimation Results....111
3.5.3 Robustness Analysis....116
3.5.4 Contributions to Income Inequality....118
3.6 Concluding Remarks....119
References....122
Appendix....136

VITA....137

List of Figures

Figure 1.1: Average Hours Worked per Week by Residents, 1976-2017, Using the ASEC Data 37
Figure 1.2: Average Hours Worked per Week, 1989-2017 37
Figure 1.3: Age-Hours Profiles by Treatment Status....38
Figure 1.4: Average Hours Worked per Week by Treatment Status 38
Figure 1.5: Changes-in-Changes Estimates 39
Figure 1.A1: Age-Hours Profiles by Cohort....46
Figure 2.1: Food Insecurity by Race/Ethnicity, 2003-2017....79
Figure 2.2: SNAP Participation Rates, 2003-2014 79
Figure 3.1: Quintile Ratio and Gini Coefficient 126
Figure 3.2: Ratios of Net Exports to GDP 126
Figure 3.3: Trends of the Related Variables of Domestic Reforms....127
Figure 3.4: Ratio of the Number of Persons Employed and Average Wage Rate Differential between Highest- and Lowest-Income Households....128

List of Table

Table 1.1: Summary Statistics 40
Table 1.2: The Impact of the Reform on Resident Hours 41
Table 1.3: Difference-in-differences Estimates of the Impact on Long-Term Labor Supply 42
Table 1.4: Estimation Results of the Impact on Long-Term Labor Supply 43
Table 1.5: The Impact of the Reform on Marriage and Divorce 44
Table 1.6: The Impact of the Reform on Fertility Decisions 45
Table 1.A1: Estimation Results of the Impact on Long-Term Labor Supply–Using Tighter Year-of-Birth Windows for Treatment and Control Groups 47
Table 1.A2: Estimation Results of the Impact on Long-Term Labor Supply–Excluding Nonrandom Selection into Treatment 48
Table 1.A3: Percentiles of the Distribution of Weekly Hours Worked 49
Table 1.A4: The Impact of the Reform on Marriage and Fertility Decisions 50
Table 1.A5: The Impact of the Reform on Marriage and Fertility Decisions 51
Table 2.1: Setup of the Numerical Example of the SNAP Decomposition 80
Table 2.2: Results of the Numerical Example of SNAP Decomposition 80
Table 2.3: Descriptive Statistics by Race and Ethnicity 81
Table 2.4: Estimated Results of the SNAP Decomposition 84
Table 3.1: Average Number of Persons Employed Per Household and Household Head Educational Attainment of First and Fifth Quintiles 129
Table 3.2: Summary of Trade Policy Changes and Domestic Reforms in Taiwan 130
Table 3.3: Unit Root Tests 131
Table 3.4: Cointegration Tests and Estimations 132
Table 3.5: Composition of Exports and Imports of Major Trade Partners (Unit: %) 133
Table 3.6: Robustness Tests 134
Table 3.7: Long-Run Mean Impacts on Income Inequality 135
Table 3.A1: Variable Definition, Source, and Summary Statistics 136

Chapter 1

The Long-Term Impact of Work-Hour Regulations on Physician Labor Supply

Judith Liu
Department of Economics and Center for Policy Research
Syracuse University, Syracuse, New York, 13244-1020
jliu61@syr.edu

1.1 Introduction

The recent expansion of the health care system and forecasts of physician shortage have made the issue of physician labor supply increasingly important. Over the last three decades, average hours worked by physicians have been falling in many developed countries, including the United States, Canada, and Australia, among others (Buske, 2004; Scott, 2006; Watson et al., 2006; Crossley et al., 2009; Staiger et al., 2010). There has also been a dramatic change in the composition of the physician workforce, with the female share of medical students rising from around 25 percent in the 1970s to around 50 percent nowadays (Chen and Chevalier, 2012). Most research and policy debates have focused on physician supply at the extensive margin (the number of practicing physicians), whereas physician supply at the intensive margin (the amount of patient care hours or services provided by practicing physicians) has been understudied (Staiger et al., 2010). A better understanding of physicians' hours of work decisions and gender differences is crucial for human resource planning purposes in the health care sector.

An important determinant of individuals' hours of work is their work experience in early career, but there is little evidence on its long-term consequences. Empirically, it is particularly difficult to identify sources of exogenous variation to test for causal effects of a transient change in labor supply. The work-hour regulations that limit the maximum hours worked by residents provide a plausibly exogenous shift in physicians' early career hours.

To become a practicing physician in the U.S., an individual must complete three to seven years of residency training after college and medical school, and then obtain medical licensure to practice medicine. Traditionally, long hours are a component of residency training, yet they may contribute to sleep deprivation which compromises patient safety. In order to reduce potential harm due to overwork of residents, the Accreditation Council for Graduate Medical Education

(ACGME) imposed regulations that restrict the average hours worked by residents to 80 hours per week and enforce standards for their duty hours in 2003. A large number of studies have examined the effects of the ACGME regulations on patients' safety and health outcomes, as well as residents' education and well-being (Philibert et al., 2013; Bolster and Rourke, 2015); yet the impacts on physicians' labor market outcomes have not been thoroughly explored in the literature. To my knowledge, the only paper that uses the ACGME regulations as a natural experiment to estimate the effects of early career hours is Wasserman (2018), which focuses on changes in specialty choice across gender.

This paper investigates whether the reform affects physicians' hours of work after they complete residency and do not face the hours constraints anymore. Using monthly data from the 1989-2017 Current Population Survey (CPS), my primary empirical strategy exploits the cohort-time variation in exposure to the ACGME regulations. As a result of the reform, the mean resident hours per week decrease by 10.03 for males and 6.87 for females. Using a difference-in-differences model with cohort and year fixed effects, the estimates suggest that exposure to the reform during residency significantly decreases mean hours worked after residency by about four hours per week, and the effects are not statistically different between male and female physicians. When taking the effects of the reform on resident hours by gender into account, a given reduction in hours during residency decreases post-residency hours significantly more for females than for males.

Since the policy limits the maximum workweek by residents, it should primarily affect those who would have worked more than 80 hours per week during residency in the absence of the regulations. To account for this disproportionate impact at the upper end of the hours distribution, I use a changes-in-changes (CIC) approach proposed by Melly and Santangelo

(2015), which estimates unconditional treatment effects of the whole distribution in the presence of covariates. Overall, the CIC estimates provide further evidence of the negative effects of the reform on long-term labor supply and show that such negative effects become stronger when moving towards the upper tail of the distribution for both male and female physicians. The greater impact among those with the longest hours confirms that the reform primarily affects the upper end of the hours distribution.

As well-documented in the literature related to physician labor supply, gender differences in hours of work may be attributable to child-rearing (e.g., Sasser, 2005; Wang and Sweetman, 2013; Wasserman, 2018), which suggests a potential mechanism of the above long-term effects. To guide our understanding about the presence of this mechanism in this context, I further examine how the reform affects male and female physicians' marriage and fertility decisions. The results show that the reform increases the probabilities of marriage and having a child, as well as the total number of children, for female physicians. In contrast, it has little impact on marriage and fertility outcomes for male physicians. These findings are consistent with previous studies and provide strong evidence on gender differences in family formation decisions in response to a policy that reduces time requirements during the prime childbearing years.

A potential mechanism for the negative impact of the reform on male physicians' long-term labor supply is through human capital accumulation. Residency can be thought of as on-the-job training to enhance physicians' skills and productivity. If physicians invest more hours during residency, they may gain more skills and have higher returns to work after residency, which increases their subsequent hours of work. The literature finds that the ACGME reform reduces continuity of care and educational continuity for residents in surgical specialties, and these losses lead to negative consequences for residents' professional development and preparedness for practice (Feanny et al., 2005; Vidyarthi et al., 2006; McBurney et al., 2008; Nakayama et al., 2009; Philibert et al., 2013). Since surgical specialties have substantially more males than females, male physicians experience a greater reduction in human capital during residency. As a result, the reform decreases their return to work rates after residency, which causes them to work fewer hours later in life. The negative impact of the reform on long-term labor supply for male physicians may be explained by this potential mechanism.

This paper has three main contributions to the existing literature. First, it exploits the work-hour regulations on residents to identify the effects of a reduction in early career hours on long-term labor supply. The findings aid our understanding of physicians' hours of work in response to the reform over time. Second, this paper explores gender differences in labor supply behaviors, along with marriage and fertility decisions as potential mechanisms. With the composition of the physician workforce changing dramatically, especially the increasing participation of females, a better understanding of how males and females respond to policy changes can suggest ways to orient policies more effectively. Third, this paper contributes to the research on dynamics of labor supply, for which it is often difficult to find a plausible identification strategy. This analysis provides important implications for broader economic theory with respect to intertemporal labor supply.

The rest of the paper is structured as follows. Section 2 provides background on the physician work-hour regulations in the United States. Section 3 discusses the conceptual framework. Section 4 describes the data and shows the effectiveness of the regulations. Section 5 presents the empirical strategy, the identification, and the mean and distributional estimated results. Section 6 addresses potential mechanisms of the effects. Section 7 concludes.

1.2 Physician Work-Hour Regulations in the United States

Medical residency training traditionally requires lengthy work hours, but there was no regulation limiting the number of hours that could be assigned to a resident physician in the United States until the late 1980s. The public and the medical education establishment started to be aware of and to investigate the consequences of overwork by residents after the death of an 18-year-old college freshman, Libby Zion, in 1984. As a result of the investigation, New York State adopted the recommendations by the committee that evaluated the training and supervision of physicians in the state, and enacted the Libby Zion Law in 1989. The law forbade residents in New York State hospitals to work more than 80 hours per week or 24 consecutive hours. It was the first regulation in the nation that restricted hours worked by residents. However, most residency programs in New York were found in violation of the law ten years after its implementation (Wasserman, 2018), and thus the law was likely not adequately enforced.

In June 2002, the Accreditation Council for Graduate Medical Education (ACGME) granted preliminary approval to similar regulations for all residents working in accredited medical training institutions in the U.S.,[1] and the regulations were implemented in July 2003 (Philibert et al., 2002). With the aim to improve patient safety by reducing fatigue-related medical errors made by residents, the ACGME's standards consist of (1) a maximum of 80 hours worked per week, averaged over one month; (2) a 24-hour limit on continuous duty with an additional six hours allowed for patient transfer, administration, and didactic lectures; (3) one day in a week free of all medically related duties; (4) a limit on call frequency; (5) a 10-hour rest period between duty periods or work shifts; (6) a maximum workweek of 88 hours allowed for

[1] All of the residency programs for doctors with a Doctor of Medicine (MD) degree and a majority of the programs for doctors with a Doctor of Osteopathic Medicine (DO) degree in the United States are ACGME-accredited. Following the proposal of the ACGME reform, the American Osteopathic Association (AOA) also adopted similar work-hour requirements.

programs in some specialties with a sound educational rationale and the approval of the Residency Review Committee. In order to comply with the policy, many residency programs changed rotation schedules, decreased call frequency, and replaced resident services with care by physician extenders (Philibert et al., 2009). With monitoring through program audits and periodic surveying of residents, penalties for non-compliance with the regulations included residency program probation and potential loss of accreditation.[2]

In 2008, the ACGME proposed minor revisions to the duty hour standards in response to the recommendations of the Institute of Medicine (IOM). The changes were made to a 16-hour limit on continuous duty for first-year residents. Residents in their second year and beyond followed the 24-hour limit with a reduction in additional hours for hand-offs from six hours to four hours. No changes were made to the 80-hour limit and call frequency. These new standards went into effect in July 2011. Though these standards were designed to improve patient safety by reducing residents' fatigue, they had also led to unintended negative consequences on residents' attainment of clinical skills. In March 2017, the ACGME further announced a policy change which raised the maximum number of consecutive hours from 16 to 24 hours for first-year residents, and this new standard went into effect in July 2017 (Asch et al., 2017).

There is an extensive literature on the effects of the ACGME regulations on residents' well-being and learning, as well as patients' safety and health outcomes. Most of the literature compares the outcomes before and after the reform in 2003 and uses observational cohort analysis from a single site, multiple sites, or national databases. Overall, the findings suggest that residents' well-being is improved between the pre- and post-2003 time periods. Many studies

[2] According to the investigation one year after the reform, five percent of the 2,235 programs that ACGME reviewed were found in violation of one of the standards. From the survey of 25,176 residents, 3.3 percent reported working more than 80 hours per week during the past four weeks.

show that residents' fatigue has decreased since the implementation of the reform (Gopal et al., 2005; Barrack et al., 2006; Hutter et al., 2006; Martini et al., 2006; Landrigan et al., 2008; Philibert et al., 2013), and some studies find that physicians are having more children, spending more time to attend family events, and leading less stressful lives since 2003 (Karamanoukian et al., 2006; Jones and Jones, 2007). However, the effects on residents' educational outcomes, patients' safety, and their health outcomes vary across studies, and some of these effects are different between medical and surgical specialties.[3]

Despite numerous studies on the reform, little is known about its effects on physicians' employment patterns in the long run. To my knowledge, the only paper directly related to labor supply effects of the ACGME regulations is Wasserman (2018), which focuses on changes in residents' specialty choice. She finds that female physicians are more likely to enter a specialty when the specialty reduces its time requirements due to the reform, but there is little change in specialty entry response among male physicians. While the regulations reduce physicians' early career hours, the question of interest is whether their long-term labor supply decisions are affected or not. The following analysis estimates the effects of the reform on physicians' post-residency hours of work, as a measure of long-term labor supply, and addresses potential mechanisms of the effects.

1.3 Conceptual Framework

To link this short-term policy to longer-term impact on labor supply, there are several distinct features of physician career paths that need to be taken into account. A physician needs to complete three to seven years of residency training after college and medical school, and then

[3] See Philibert et al. (2013) and Bolster and Rourke (2015) for a systematic literature review.

obtain medical licensure to practice medicine. Residency programs are typically run by hospitals and have a limited number of residency slots each year. Residents are paid on an annual basis, which is largely funded by the government, and there is little difference in resident salary across specialties. Since residents are forced to work a set of hours, their labor supply can be considered as perfectly inelastic within specialties. Therefore, physicians are likely to have less leisure time during residency because they are constrained to work more hours than they would otherwise choose. Despite the number of hours worked during residency, their average hours typically decrease after completing the training. The wage increases dramatically after residency as most practicing physicians make considerably more than resident salary, and post-residency salary varies largely across specialties.

Since residents' hours of work are constrained by institutional rules regulating labor time and effort provision, the intertemporal substitution hypothesis with time separable utility does not fit in this context. Alternatively, a theoretical hypothesis that can be used to explain the long-term labor supply effect is the neoclassical model with non-separable utility (Fehr and Goette, 2007). Holding the wage constant, this model predicts that an increase in a worker's effort in the previous period causes a higher disutility of labor in the following period, which decreases the worker's labor supply. Since the ACGME reform causes an anticipated transitory reduction in physicians' labor supply in early career without changing their wage and lifetime wealth, it decreases their disutility of effort during residency. Subsequently, it will increase their labor supply after residency, based on the prediction of this model.

An opposite theoretical hypothesis is the "persistence hypothesis" which states that individuals' work experience in early career is a major determinant of subsequent labor supply due to human capital accumulation, change in family commitments, and taste for work, among others (Clark and Summers, 1982). On human capital grounds, those who work more tend to accumulate more human capital, which in turn increases the return to work relative to leisure in the future (Heckman and Willis, 1979; Freeman, 1980; Clark and Summers, 1982). Those with a lack of work experience, on the other hand, may develop family commitments which reduce the return to work relative to staying outside of the labor force (e.g., Mincer and Polachek, 1974; Polachek, 1975; Becker, 1985, Gronau, 1988; Angrist and Evans, 1998; Sasser, 2005; Goldin, 2014; Kleven et al., 2018). In addition, individuals' taste for work could be affected by prior work experience according to habit formation effects (Clark and Summers, 1982; Clark, 1999).

This hypothesis suggests that a short-term reduction in physicians' early career hours tends to persist after residency. Intuitively, human capital accumulated through residency experience affects labor supply in the future. If physicians invest more hours during residency, they may gain more skills and have higher returns to work after residency, which increases their subsequent hours of work. The literature finds that the work-hour limits reduce continuity of care and educational continuity for residents (Feanny et al., 2005; Vidyarthi et al., 2006; McBurney et al., 2008; Nakayama et al., 2009), and these losses lead to negative consequences for residents' professional development and preparedness for practice, especially in surgical specialties (Philibert et al., 2013). Since there are substantially more male physicians in surgical specialties, they experience the greatest reduction in resident hours as well as human capital accumulation. As a result, the reform decreases their return to work rates after residency, which causes them to work fewer hours later in life.

Other potential mechanisms pertain to family formation decisions. Since the work-hour regulations affect residency training, which occurs during the prime childbearing years, physicians would plan the timing of marriage and fertility relative to their residency.

Theoretically, the regulations alter the labor market conditions for residents and thus change their opportunity cost of time factored into the fertility transition. Previous research has shown that the reform results in physicians having more children, spending more time to attend family events, and leading less stressful lives (Karamanoukian et al., 2006; Jones and Jones, 2007). Consequently, these changes in family commitments may decrease their labor-force attachment and keep them from developing further their careers.

The dynamic impact of children on labor market outcomes also greatly depends on spousal income (Goldin, 2014). In two-income households, if their partner is doing well financially, they may feel more comfortable pulling back on their hours. As such, physicians with higher-earnings spouses have a lower opportunity cost of career interruptions (Sarma et al., 2011). According to the AMA Masterfile, nearly 40 percent of physicians marry another physician or health care professional. In addition, most of the female physicians are married to male physicians, while the reverse is not true (Sasser, 2005). With higher-earnings physician spouses, who also work long hours, new physician mothers face more binding constraints on hours and a lower opportunity cost of career interruptions. Therefore, they are more likely to reduce their post-residency hours than male physicians, who are less likely to have physician spouses. In Section 6, I provide empirical evidence on the mechanisms pertaining to marriage and fertility across gender.

Overall, the neoclassical model with non-separable utility predicts that the ACGME regulations decrease physicians' disutility of work during residency and thus increase their labor supply in the long run. In contrast, the persistence hypothesis suggests that the regulations reduce resident hours and at the same time lower the opportunity cost of work time, especially for those who would have worked more than the work-hour limits in the absence of the reform. This could lead to less human capital, more family commitments, and habit formation effects for those exposed to the reform, and thus results in a reduction in hours worked over time. From the above discussion, the potential impact of the ACGME regulations on long-term labor supply is ambiguous due to the contradicting effects between these two hypotheses; therefore, empirical evidence is needed to better understand physicians' employment patterns. As it will be presented in Section 5, my empirical results are consistent with the persistence hypothesis and suggest that the short-run reduction in labor supplied persists even when physicians are not bound by the work-hour limits.

1.4 Data and the Effectiveness of the Reform

1.4.1 Data Construction and Summary Statistics

This analysis uses data from the monthly Current Population Survey (CPS) between 1989 and 2017. Administered by the U.S. Census Bureau, the monthly CPS is a household-based survey which selects a nationally representative sample and contains a large amount of demographic and employment information. To identify physicians in the CPS, I restrict the sample to the civilian non-institutional population who hold an advanced degree and reported their occupation as a "physician or surgeon."

Whether a physician was exposed to the work-hour regulations is based on the year of residency training, but such information is not available in the CPS. Inspired by Staiger et al. (2010), I identify physicians as residents if they were younger than 35 and use the year of birth as a proxy for exposure to the ACGME regulations.[4] Physicians who could have been potentially

[4] According to the 2007 AMA Physician Masterfile data, which is the primary source of physician workforce data in the U.S., Staiger et al. (2010) point out that 97% of hospital-based physicians younger than 35 were residents. However, not all residents were trained in a hospital-based program, and thus using age 35 to identify residency status might lead to a potential source of bias. This problem is addressed in Section 5.3.

subjected to the regulations were born after 1968 (i.e., those who worked as a resident after 2003), and they are categorized as the treatment group. For physicians trained in New York, although they might have been potentially exposed to similar regulations, the Libby Zion Law, most residency programs in New York were found in violation of the law (Wasserman, 2018). In addition, I use the CPS Annual Social and Economic Supplement (ASEC) data to test whether there was a significant decrease in resident hours around the implementation of the law in 1989.[5] Figure 1.1 shows that the average hours worked by residents remained fairly stable around 1989 and changed little through the 1980s and 1990s, suggesting that the law was inadequately enforced.[6] On the other hand, the average hours worked by residents decreased sharply following the imposition of the work-hour limits in 2003, demonstrating that the ACGME regulations effectively led to hours cut. Therefore, the central variation in the empirical analysis below comes from the cohort-time variation in exposure to the 2003 ACGME reform.

According to the U.S. Department of Health and Human Services, more than 95 percent of the physicians graduated from medical school after age 26. In addition, by the end of the sample period in 2017, the oldest possible age that the treatment group can achieve is 48. Therefore, the analysis focuses on physicians aged 26 to 48 with non-missing values for weekly hours worked.[7] The analysis sample comprises 70,868 physicians. It is worth noting that this

[5] The monthly CPS does not provide the hours worked variable before 1989, which is the main reason why the analysis period starts from 1989.

[6] I also look at the trend for New York only and find no significant change in hours worked by residents around 1989 either. However, the sample size of New York physicians in the ASEC is very small (around 30 observations per year on average), and thus it is not informative enough.

[7] The measure of hours worked is based on the self-reported hours in the previous week in the monthly CPS. An alternative measure is the usual number of hours per week (over an unspecified time period), but it is not available until 1994 in the monthly data. I chose the former for the analysis since it has a relatively shorter-term recall and is available for a longer period of time. Note that the hours worked measure is top-coded at 99 hours prior to 1994; however, there are only 690 observations (about 1%) in the sample at the top-coded value of 99 before 1994.

selected group are at their prime working and childbearing ages, which helps understand the role of job flexibility in the work-family interface.

Table 1.1 shows the summary statistics of this analysis sample. The mean age of the sample is 37.61, and approximately 66 percent of the subjects are male. The average hours worked per week is 54.62 (standard deviation = 18.37). With respect to gender differences, male physicians tend to be slightly older and have smaller proportion of the treated population due to the increasing female share of physicians in recent decades. In addition, male physicians work about seven hours more than their female counterparts, and they have a higher rate of marriage and have more children on average. With respect to differences by treatment status, the treatment group is older and comprises more females. The average hours worked per week is 54.40 for the treatment group and 54.89 for the control group.

There are many advantages of using the monthly CPS data for this analysis. First, it provides repeated cross-sectional observations over a longer time period than any other comparably sized dataset that includes physicians' information. This is particularly useful for analyzing the effects on lifecycle patterns. The large enough sample also helps conduct analyses separately by subpopulations and run robustness tests using different regression specifications. Second, the CPS includes important demographic characteristics and employment information. These variables matter for identification because they allow us to account for dynamic changes in the composition of the physician workforce that might cause potential imbalances in the demographic characteristics between the treatment and control groups. Third, the CPS data are more up-to-date than the American Medical Association (AMA) Physician Masterfile in terms of physicians' employment information and can be used as a benchmark dataset on physician labor supply (Staiger et al., 2009). Fourth, it is widely conjectured that residents' self-reported hours from the ACGME monitoring data may underestimate hours worked due to the desire to protect residency programs or pressure from residency program directors (Landrigan et al. 2006; Szymczak et al. 2010; Fargen and Rosen, 2013). The CPS data, collected by non-ACGME researchers, limit the potential for this type of misreporting.

Nevertheless, the CPS lacks physician-specific information regarding residency training and specialty choice. Therefore, I cannot directly identify the treatment status and analyze some of the other interesting labor market outcomes. In addition, the information on income is top-coded in the CPS for confidentiality reasons. More than 85 percent of the sample analyzed here has top-coded values or missing values in weekly earnings at ages 35 to 48.[8] For these reasons, it is difficult to provide direct evidence on how the regulations change physicians' earnings profiles over time.

1.4.2 First Stage: The Effectiveness of the Reform

To assess the effectiveness of the regulations for the analysis, I plot the trends in hours worked by resident and non-resident physicians using the analysis sample in Figure 1.2. Prior to 2003, residents were not exposed to the work-hour limits. As shown, the average hours worked per week by residents remained high through 2002 and then declined sharply after the preliminary approval of the ACGME reform in 2002 and its implementation in 2003. The average resident hours per week decreased from 63 in 2002 to 58 in 2004. This sharp decline after the introduction of the reform provides evidence that the regulations were enforced.

On the contrary, such plummet was not found in the work hours trend of nonresident physicians since they were not restricted by the regulations. Instead, their hours worked have

---

[8] Total weekly earnings are top-coded at $1923 prior to 1997 and $2885 since 1998. For hourly wages, more than 90 percent of the physicians in the CPS have missing values.

gradually trended downward since the 1990s. This trend can be largely explained by the aging of the physician population and the increasing proportion of female physicians, who practice fewer hours than their male counterparts on average (Crossley et al., 2009; Staiger et al., 2010; Sarma et al., 2011, Wang and Sweetman, 2013). Staiger et al. (2010) also point out other possible factors that drive the downward trend of hours worked by practicing physicians, such as the decrease in physician fees since the early 1990s, developments among both public and private payers in the 1990s, and improvements in physician productivity due to technology. In addition, the rapid growth in care by hospitalists in the late 1990s and the early 2000s (Kuo et al., 2009) and the HMO penetrations in the 1990s (Zhan et al., 2004) may also be attributed to the decreasing hours worked by practicing physicians.

Figure 1.3 compares the trends in hours worked between physicians who were exposed to the regulations and those who were not. As a result of the reform, the treatment group (physicians born after 1968) worked fewer hours per week than the control group (physicians born before 1968) during residency (ages 26 to 34). Interestingly, the difference in hours worked between the treatment and control groups still remained even when physicians were not constrained by the work-hour limits after residency (above age 35), showing a persistent decline in hours worked. Compared to physicians, there is no persistent and significant difference in hours worked between the younger and older cohorts for the other professions that also make a fairly high salary and require an advanced degree (e.g., lawyers and dentists), as shown in Figure 1.A1 in the Appendix. This comparison provides additional evidence on the substantial impact of the reform on physician labor supply.

In addition to visual evidence, I estimate the effect of the reform on resident hours for the full sample and by gender. Since residency programs must comply with all the work-hour standards, they should be considered as a whole when interpreting the estimated effects. The empirical specification is as follows:

$$\mathrm{Y}_i = \beta_0 + \beta_1 \mathrm{Exposed}_i + X_i\beta_2 + \alpha_c + \alpha_t + \alpha_s + \varepsilon_i, \tag{1}$$

where Y_i is the weekly hours worked by physician i aged 26 to 34. $\mathrm{Exposed}_i$ is a dummy variable indicating exposure to the ACGME regulations during residency training, which equals 1 for physicians born after 1968 and 0 for physicians born before 1968. X_i is a set of covariates including age, gender, and race/ethnicity. α_c, α_t, and α_s denote cohort, year, and state fixed effects, respectively. The estimates of β_1 identify the effects of the reform on resident hours and are shown in the first row of Table 1.2. As a result of the reform, the mean resident hours per week significantly decreased by 8.38 overall, which supports the effectiveness of the ACGME reform. Analyzing by gender, the reform reduced the mean resident hours per week by 10.03 for males and 6.87 for females.[9]

1.5 The Impact on Long-Term Labor Supply

1.5.1 Difference-in-Differences Approach

A fundamental challenge in interpreting the pattern shown in Figure 1.3 as causal is that the cohort variation that identifies differences in exposure to the regulations is time-series in nature. Omitted variables that are correlated with changes in labor supply and the exposure, and

[9] This effect is smaller among females than among males for two reasons. First, there are fewer female physicians whose hours were capped by the regulations. In my analysis sample, only the top 20 percent of the female control group exceeds 80 hours worked per week. Wasserman (2018) also shows that females were less likely to choose the most time-intensive specialties where the hours worked by residents were more than 80 hours per week before the reform. Hence, the majority of the female physicians were not primarily affected by the reform, and the mean estimate of the effect on resident hours is attenuated. Second, females are found to enter more time-intensive specialties as a result of the reform, whereas there is little change in males' specialty choice (Wasserman, 2018). If females change their specialty choice in response to the reform, they are also potentially altering their residency hours towards longer hours. This behavioral change increases mean hours worked by female residents and balances out the effect on resident hours which were originally designed to reduce hours worked.

other secular trends that affect physician hours, might explain this pattern as well, resulting in an identification problem. To tackle this problem and identify the causal impact, I begin by the baseline difference-in-differences approach using the cohort-time variation in exposure to the ACGME regulations. Since the reform affects residents trained after 2003, the strategy is to compare the change in hours worked from residency to post-residency between the treatment and control groups. The regression framework of the baseline model is as follows:

$$\mathrm{Y}_{ict} = \beta_0 + \delta_1 \mathrm{Exposed}_c + \delta_2 \mathrm{Post}_t + \delta_3 (\mathrm{Exposed}_c \cdot \mathrm{Post}_t) + \varepsilon_{ict}, \tag{2}$$

where *i* indexes physicians, *c* indexes birth cohorts, and *t* indexes years. The outcome variable Y_{ict} is the weekly hours worked by physicians. Similar to the definition in Equation (1), $\mathrm{Exposed}_c$ is an indicator of exposure to the ACGME regulations during residency training. Post_t is an indicator of completing residency in year *t*, which equals 1 for physicians aged 35-48 and 0 for physicians aged 26-34. δ_3 is the coefficient of interest. To identify it as the casual impact, the treatment and control groups are assumed to have the same trends in hours worked over time in the absence of the regulations.

Figure 1.4 shows the trends in post-residency hours between the treatment and control groups after the implementation of the 2003 ACGME reform. There does not seem to be a significant difference between groups over time. However, it is unclear from this figure regarding the role of the reform since there are some factors that affect labor supply and also correlate with exposure to the regulations, as suggested by the summary statistics in Table 1.1. In particular, the control group consists of older physicians, who are likely to work fewer hours per week, whereas the treatment group consists of more female physicians, who are likely to work less than their male counterparts.

To account for potential imbalances in the demographic characteristics between the treatment and control groups, I control for pre-treatment baseline observables. I also expand the model by including cohort and year fixed effects to control for additional unobserved factors. Compared to the inclusion of the two indicators, Exposed_c and Post_t, these fixed effects flexibly span the cohort-time variation. The empirical specification can be written as:

$$\text{Y}_{ict} = \beta_0 + \alpha_c + \alpha_t + \beta_1(\text{Exposed}_c \cdot \text{Post}_t) + X_{ict}\beta_2 + \alpha_s + \varepsilon_{ict}. \tag{3}$$

X_{ict} is a set of pre-determined demographic controls including age, gender, and race/ethnicity. α_c reflects fixed effects for birth cohorts, and α_t reflects fixed effects for the calendar year in which they are observed. The cohort fixed effects control for differences across cohorts in the outcome variable, and the year fixed effects control for any general time trends in the outcome variable, picking up some of those possible factors that drive the long-term secular decline in physician work hours mentioned in Section 4.2. I also add a set of state dummies α_s to control for any time-invariant unobservables that affect the outcome variable across states. In particular, it accounts for the state differences in the institutional design features, such as state-specific licensing requirements. I do not include controls which may cause potential endogeneity with respect to labor supply (e.g., family formation and practice setting). The coefficient β_1 identifies the treatment effect by contrasting the hours worked from residency to post-residency between physicians who were exposed to the regulations and those who were not.

Previous studies have documented differences in labor market outcomes between male and female physicians (e.g., Rizzo and Blumenthal, 1994; Sasser, 2005; Rizzo and Zeckhauser, 2007; Wang and Sweetman, 2013; Wasserman, 2018). In addition, my regression results (shown in Section 5.2) also indicate that gender has a significant impact on long-term labor supply. In addition to the analysis on the full sample, I also estimate the effects by gender to explore whether the ACGME regulations affect male and female physicians differently.

The key identifying assumption in Equation (3) is the so-called parallel trends assumption, that is, the evolution of hours worked between the treatment and control groups (conditional on observed variables) is the same over time in the absence of the reform. This analysis has several advantages in meeting this condition. First, the ACGME regulations were enacted in order to reduce fatigue-related medical errors made by residents. The motivation and the nature of this policy make it unlikely to be correlated with other policies that affect physician labor supply or costs of childrearing. Second, the sample is fairly homogeneous with respect to skills and other traits. Since this analysis is based on a profession that has highly competitive entry requirements, rigorous educational standards, and very specialized training, the general concern about unobserved heterogeneity across individuals or cohorts is considerably diminished. Since the model includes cohort and year fixed effects, and a set of demographic controls, the estimated effects on the outcome variables can be attributed to changes in hours worked during residency training within cohorts over time.

1.5.2 Estimated Mean Effects

Table 1.3 shows the estimation results using the baseline difference-in-differences model, specified in Equation (2), for the overall sample and by gender. Each cell contains the mean hours worked for its subgroup of the sample. For the overall sample shown in Panel A, the residency versus post-residency difference in hours is 9.07 for the treatment group and 7.15 for the control group. Thus, the treatment group worked 1.93 less hours per week than the control group from residency to post-residency. With a standard error of 0.87, it is statistically different from zero at the 5 percent level. The analysis by gender is shown in Panels B and C. Among male physicians, the treatment group decreases weekly hours worked by 1.09 (or 2.26 percent) more than the control group, but the estimate is statistically insignificant. Among female physicians, the treatment group decreases weekly hours worked by 1.63 (or 2.90 percent) more than the control group.

Controlling for observed demographic differences between groups and the cohort and year fixed effects, the preferred estimates of the long-term labor supply effects using Equation (3) are shown in Table 1.4. Columns (1) and (2) display the results for the full sample, without and with the inclusion of demographic controls, respectively. The first row reports the estimated coefficients of interest. The estimate in Column (2) implies that the reform decreases physicians' mean hours worked at ages 35 to 48 by 4.28 hours per week, which is a statistically significant 7.99 percent decrease over the control group mean of 53.59 hours. Columns (3) to (6) show the estimates by gender. The results indicate that the regulations reduce both male and female physicians' long-term hours worked by about four hours per week on average, and the effects are not statistically different across gender (p-value = 0.67). These findings are robust to the inclusion of the covariates. The next six rows in Columns (2), (4), and (6) report the estimated coefficients of the demographic characteristics on long-term hours worked. Among all three columns, the average hours worked decrease significantly with age, and as shown in Column (2), gender seems to have a large and significant impact on post-residency hours, with males working more than females by 7.09 hours per week.

Since men tend to enter specialties that require longer hours worked, there are more male physicians whose hours were capped by the regulations. To obtain the effects of a given reduction in hours during residency on subsequent labor supply, I take into account the effects of the reform on resident hours and divide the reduced form coefficients in Columns (2), (4), and (6) by their corresponding first-stage estimates. For the full sample, a reduction in hours during residency decreases labor supply at ages 35 to 48 by 0.51 hours per week on average (standard error = 0.02). Analyzing by gender, a reduction in hours during residency results in a decrease of 0.39 hours (standard error = 0.02) for males and 0.65 hours (standard error = 0.05) for females after residency.[10] Although female physicians in general were less likely to be restricted by the work-hour limits, these findings suggest that females are more responsive to a given reduction in early career hours caused by the reform, and this gender difference is statistically significant at the 1 percent level.

There are a few observations with likely unreliable self-reported hours in the data.[11] To address this concern and reduce the impact of possibly spurious outliers, I repeat the analysis using winsorized and trimmed hours worked at the 1% and 99% levels, as well as the 5% and 95% levels, as alternative outcome variables. Winsorizing at the 1% and 99% levels sets the bottom 1% to the 1th percentile and the top 1% to the 99th percentile. Without discarding the extreme values, the winsorization method still takes those values into account and treats them as if respondents reveal certain information on their hours worked. The estimated effects using these alternatives are very similar to the results shown above.[12]

1.5.3 Potential Threats to Identification

Although the inclusion of control variables and the homogeneity of the physician workforce ameliorate potential threats from unobserved confounders, there may still be at least

[10] These estimates are obtained by -4.28/-8.38 = 0.51 for the full sample, -3.96/-10.03 = 0.39 for males, and -4.45/-6.87 = 0.65 for females.
[11] The maximum of self-reported hours worked is 192 in the data, which is obviously exaggerated.
[12] The results are available upon request.

two other concerns regarding the identification. The first problem pertains to the year-of-birth proxy for treatment status. Ideally, I would use the year of residency to identify individual exposure to the ACGME reform, but the CPS does not contain such information. This may result in misclassifications of the treatment and control groups. The second problem pertains to any remaining unobserved heterogeneity of exposure to the regulations with respect to labor supply. I discuss below these identification issues and how I attempt to assess them.

First, using age 35 to identify residency status is a potential source of bias. Depending on the medical specialty, the length of residency training ranges from three to seven years.[13] This leads to some variation in the age that a physician can complete residency.[14] There are two possible misclassifications of the treatment status: (1) physicians who were born after 1968 and completed residency training before 2003, and (2) physicians who were born before 1968 and completed residency training after 2003. In the first case, the treatment group contains non-treated individuals; in the second case, the control group contains treated individuals. As a consequence, the magnitude of the estimated effects would be underestimated in both cases.

To assess the effects of potential misclassifications, I estimate three alternative specifications which reclassify the treatment and control groups with tighter year-of-birth windows, and the age proxies for residency status are also adjusted accordingly. Since the main analysis uses physicians born before and after 1968 as the control and treatment groups, respectively, the alternative specifications adjust the year-of-birth proxies for the treatment status as follows: (1) the treatment (control) group consists of physicians born after 1969 (before 1967);

[13] For example, internal medicine, general surgery, and neurosurgery require three, five, and seven years of training, respectively.

[14] Nearly all physicians graduate from medical school after age 26. With a minimum of three years for residency training, the earliest possible age to complete residency is 29. Similarly, with a maximum of seven years for residency training and allowing a gap of five years at some point, the oldest possible age to complete residency is likely to be 38.

(2) the treatment (control) group consists of physicians born after 1970 (before 1966); (3) the treatment (control) group consists of physicians born after 1971 (before 1965). When using tighter year-of-birth windows for the treatment status, there should be less misclassification; however, it discards non-negligible amount of observations. As shown in Table 1.A1 in the Appendix, the results using these alternative treatment and control groups are similar to the main results.

Second, there may still exist some unobserved factors that cause nonrandom selection of individuals into the physician workforce following the implementation of the reform. It could be that the reform changed the pool of applicants and entrants into the medical profession in dimensions not captured by the admission criteria and observed characteristics, but are relevant to the labor market. Specifically, given the decline in hours requirements during residency, individuals who prefer balanced lifestyles would be induced to enroll in medical school. These unobserved preferences are correlated with a priori disposition toward fewer hours worked at later ages, which would lead to a decrease in the average hours worked by physicians over time. As a result, the magnitude of the negative effects of the reform on long-term labor supply would be overstated.

To assess whether the effects are driven by this selection bias, I estimate an alternative specification, taking as the treatment group physicians who already entered medical school at the time of the reform but were trained under the new regulations during residency. Given that the ACGME regulations were approved in 2002 and that the fresh college graduates in 2002 were born in 1979-1980, I restrict my sample to the 1941-1980 cohorts and re-estimate the effects. Since there may be some physicians enrolling in medical school few years after college, I also restrict the sample to the 1941-1978 cohorts and the 1941-1976 cohorts to test the robustness of my results. As shown in Table 1.A2 in the Appendix, the key estimates are robust to these alternative specifications, suggesting that the likelihood of this self-selection confounding the results is not considerable.

Although unobserved heterogeneity seems less applicable to physicians who have been through highly competitive admission process and invested many years in formal training beyond college with the completion of medical school and residency training, the above robustness check suggests that remaining unobservable heterogeneity is not a significant concern. Other empirical studies also find no evidence of reduced hours worked driven by unobserved preferences for balanced lifestyles among younger physician cohorts (Crossley et al., 2009; Staiger et al., 2010; Sarma et al., 2011).

1.5.4 Distributional Effects

According to the ACGME work-hour standards, the regulations should primarily affect residents who would have worked more than the work-hour limits (e.g., 80 hours per week) in the absence of the reform. This naturally leads to a disproportionate impact on those at the upper tail of the hours distribution. As shown in Table 1.A3 in the Appendix, only those above the 75th percentile among the control group aged 26-34 (unaffected residents) work over 80 hours per week. The estimated mean impacts may be attenuated by the other part of the distribution and incompletely reveal the effects on those affected.

To identify the heterogeneous impacts across the hours distribution, I use a changes-in-changes (CIC) model following Athey and Imbens (2006) and Melly and Santangelo (2015) that extends the model to include covariates. The CIC framework relaxes the parallel trends assumption and provides unconditional treatment effects of the whole distribution. The estimated quantile treatment effects provide evidence on what would happen to the overall hours distribution in the long run if there is a policy regulating the maximum workweek during residency. I estimate the CIC effects for 17 quantile values, $q = \{0.1, 0.15, \ldots, 0.85, 0.9\}$, and their bootstrapped 95-percent confidence intervals with 1,000 replications, controlling for age, gender, race/ethnicity, and cohort and year fixed effects.

Figure 1.5 shows the CIC estimates for the full sample (Panel A) and by gender (Panels B and C). In general, the effects become negative and stronger when moving towards the upper end of the work hours distribution. As shown in all three panels, the quantile treatment effects are not statistically different from zero between the 10^{th} and the 70^{th} quantiles, and the effects become negative and larger than the mean estimates at the top of the distribution. This finding confirms that the reform primarily affects those with the longest hours of work. For the full sample, the estimates above the 80^{th} quantile are statistically different from zero and remain stable at around -5.36. For male physicians, the estimates above the 75^{th} quantile are statistically significant and have an average level of -8.47. For female physicians, the estimates are imprecisely estimated, but overall negative and hovering around -5.11 for those above the 70^{th} quantile (except for the 80^{th} quantile). At the extreme upper tail of the distribution (the 90^{th} quantile), the magnitudes of the effects for male and female physicians are similar.

There are two potential reasons why the effects are greater for males than for females at most of the upper quantiles. First, since hours worked by male physicians at the upper quantiles are more likely to be capped by the regulations, the reform reduces their long-term labor supply more than their female counterparts. Second, more female physicians enter the most time-intensive specialties as a result of the reform, whereas there is little change in specialty choice by male physicians (Wasserman, 2018). If more female physicians enter long-hours specialties in response to the reform, they are also potentially altering their residency hours and increasing hours worked after residency. The link between changes in specialty choice and hours worked provides a potential explanation and mechanism for the gender difference in the effects of the reform on the distribution of hours worked.[15]

1.6 Mechanisms

The work-hour regulations reduce physicians' early career hours and at the same time lower the opportunity cost of work time. As discussed in Section 3, this could lead to less human capital, more family commitments, and change in taste for work for the treatment group. Therefore, the reform results in a negative impact on physicians' long-term labor supply. Since the regulations affect residency training, which occurs during the prime childbearing years, the mechanisms pertaining to marriage and fertility decisions are particularly of interest. Various studies suggest that home production is disproportionately undertaken by females even within this highly skilled profession (e.g., Sasser, 2005; Wang and Sweetman, 2013; Wasserman, 2018). Women may choose the specialty and work environment that are family friendly, and avoid jobs with long hours and greater career advancement possibilities. Wang and Sweetman (2013) show that married female physicians work fewer hours per week than both their married male counterparts and their unmarried female counterparts. The impact of children on women's labor market outcomes is large and persistent, whereas there is little evidence on such impact on men.

The CPS data allow us to learn about the presence of marriage and fertility mechanisms in the context of the ACGME reform and whether there are gender differences in these

[15] There are two caveats of these estimated quantile treatment effects. First, the existence of point masses in the hours worked data might contaminate the effects, and thus the results must be evaluated with caution. Second, whether men or women are more responsive to a given reduction in resident hours across the hours distribution is unknown without knowing the corresponding distributional effects of the reform on resident hours.

mechanisms. I use the following regression framework to investigate the impact of exposure to the regulations on male and female physicians' marriage and fertility outcomes:

$$\mathrm{Y}_i = \beta_0 + \beta_1 \text{Exposed}_i + X_i\beta_2 + \alpha_t + \alpha_s + \varepsilon_i.^{16} \tag{4}$$

When examining the marriage mechanism, I use binary indicators of marriage and divorce for physicians between ages 26 and 48 as the outcome variables. For the effects on fertility decisions, I examine two outcomes: (1) fertility at the extensive margin, which is an indicator of having at least one child; (2) fertility at the intensive margin, which is the total number of children (completed fertility).[17] By contrasting these outcomes between physicians who were exposed to the regulations and those who were not, the estimates of β_1 show the effects of interest. The identification is based on the selection-on-observables assumption; that is, there is no unobserved factor that affects both outcomes (marriage and fertility) and treatment (exposure to the reform). This assumption plausibly holds since the reform is not correlated with other policies that would affect physicians' marriage and childrearing. The set of control variables and the homogeneity of the physician workforce also make the identification assumption more plausible to be satisfied.

Table 1.5 shows the effects of the work-hour regulations on the likelihoods of marriage and divorce between ages 26 and 48, controlling for demographic characteristics (age, gender, and race/ethnicity), year fixed effects, and state fixed effects. Columns (1) and (2) show the results using the full sample, and Columns (3) to (6) present the effects by gender. The estimated

[16] Note that this is not a difference-in-differences model. Contrary to hours of work, there is little variability in marriage and fertility outcomes during residency, and most of the variability comes from the post-residency period. Instead of using a difference-in-differences model which contrasts the outcomes from residency to post-residency between groups, I use the regression framework specified in Equation (4) to capture the overall effect of the reform on marriage and fertility outcomes.

[17] According to the National Association for Public Health Statistics and Information Systems, completed fertility is defined as the number of children to a person by the end of a woman's childbearing years, 15 to 44 years old, the latest age at which people typically have their last child.

coefficients of interest, shown in the first row, suggest that the reform significantly increases the probability of marriage by 9.2 percentage points and slightly decreases the probability of divorce by 1.5 percentage points for female physicians. Conversely, there is no significant impact on male physicians' marriage and divorce rates. These estimates suggest that marriage decisions made by females are more sensitive to the work-hour regulations than those made by males.

Table 1.6 presents the estimated effects of the regulations on physicians' fertility between ages 26 and 48. Columns (1), (3), and (5) show the effects on fertility at the extensive margin for the full sample and by gender. The estimates suggest that the reform increases the probability of having a child for female physicians by 9 percentage points, which is statistically significant at the 5 percent level. However, the reform does not affect the probability of having a child for male physicians. Columns (2), (4), and (6) show the estimated effects on completed fertility. The results also suggest a substantial gender difference. Exposure to the regulations leads to a significant increase of 0.2 children for female physicians, but it does not affect the total fertility of male physicians.[18] In addition, I estimate the effects for physicians aged 26-34 and 35-48 separately, as shown in Table 1.A5 in the Appendix. These estimates provide suggestive evidence on how the reform changes the timing of marriage and fertility beyond the total impact of the reform. The results indicate that the reform increases the marriage and fertility outcomes in the post-residency period more than those during residency, but they are not statistically different from each other.

To sum up, I find substantial gender differences in both marriage and fertility decisions in response to the reform. The regulations raise the likelihood of marriage and have positive and significant effects on fertility at both the extensive and intensive margins for female physicians.

[18] The estimation results using probit and logit models for the binary outcomes of marriage and fertility are similar to the above results, as shown in Table A4 in the Appendix.

On the contrary, the regulations have little impact on marriage and fertility outcomes for male physicians. During the childbearing years, these effects for females may result in a potentially supply shift that accounts for at least some of the decrease in their long-term labor supply. As discussed in Section 3, most female physicians are married to physician spouses, who are likely to have high earnings and work long hours. As the marriage and fertility rates increase for female physicians due to the reform, they likely face greater household obligations and more binding constraints on hours with a lower opportunity cost of career interruptions. As a result, this situation would lead them to work fewer hours, more regular schedules, and generally more conducive to combining career and family. The estimated results in this section provide empirical evidence to support potential mechanisms of females' labor supply responses to a policy that reduces time requirements during the prime childbearing years. Changes in family commitments likely account for females' greater responsiveness to a given reduction in hours. These findings are also consistent with previous studies on gender differences in the relationship between childbearing and labor market outcomes (e.g., Mincer and Polachek, 1974; Polachek, 1975; Becker, 1985, Gronau, 1988; Angrist and Evans, 1998; Sasser, 2005; Wang and Sweetman, 2013; Goldin, 2014; Kleven et al., 2018; Wasserman, 2018).

1.7 Conclusions

This paper estimates the impact of the work-hour regulations that limit the maximum hours worked by residents on physicians' long-term labor supply. As a result of the 2003 ACGME reform, the mean resident hours per week significantly decrease by 10.03 for males and 6.87 for females. Exploiting the cohort-time variation in exposure to the reform, I contrast the hours worked from residency to post-residency between the treatment and control groups using a difference-in-differences approach. The results suggest that the reform significantly reduces mean hours worked after residency by about four hours per week for both male and female physicians. When taking the effects of the reform on resident hours into consideration, women seem to be more responsive to a given reduction in early career hours caused by the reform.

The heterogeneous impacts across the hours distribution are revealed by a changes-in-changes model. The estimated quantile treatment effects show that the reform does not have a statistically significant impact on those below the 70th quantile of the hours distribution. When moving towards the upper tail of the distribution, the effects become significantly negative and larger than the mean estimates, which confirms that those at the upper end of the distribution are primarily affected by the reform. Since hours worked by male physicians are more likely to be capped by the regulations, the reform reduces their long-term labor supply more than their female counterparts at the upper quantiles. However, at the extreme upper tail of the distribution, the magnitudes of the effects for male and female physicians are similar.

To reveal potential mechanisms of the effects uncovered on long-term labor supply, I examine how the regulations affect physicians' marriage and fertility outcomes across gender. The empirical evidence suggests substantial gender differences in marriage and fertility choices in response to the reform that reduces work time requirements during the prime childbearing years. It indicates that changes in family commitments could be potential mechanisms for females' long-term labor supply effects. On the other hand, since there are substantially more male physicians in surgical specialties that suffer the greatest reduction in hours due to the reform, they may experience detrimental effects on their professional development and preparedness for practice. The mechanism of human capital accumulation might potentially account for males' long-term labor supply effects, although I am not able to explore this mechanism empirically.

With increasing participation of females in the physician workforce and limited evidence on long-term labor supply responses to a reduction in early career hours, the physicians' hours of work decisions and potential gender differences are substantive issues for human resource policies in the health care sector. With respect to the effects of the reform on long-term labor supply, the reduction of four hours per week among practicing physicians in younger cohorts does not seem small. Policy makers may need to take into account changes in the amount of patient care hours or services provided by practicing physicians when addressing the projected supply of healthcare. With respect to gender differences in marriage and fertility decisions, the results indicate that less time requirements may help women plan the timing of marriage and fertility relative to their residency. Since residency training occurs during the prime childbearing years, and nowadays almost half of the medical students are women, developing workplace policies to accommodate pregnancy and childbearing is important for the medical profession. Future research may consider using other physician surveys that include information on earnings and other dimensions of physicians' labor market outcomes. In particular, dynamic changes in earnings profiles caused by the regulations can give important insight into labor supply behavior. The findings can also provide direct evidence of the resulting effects on the gender wage gap.

References

Angrist, J. D. and Evans, W. N. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review* 88 (3): 450-477.

Asch, D. A., Bilimoria, K. Y., and Desai, S. V. 2017. "Resident Duty Hours and Medical Education Policy – Raising the Evidence Bar." *New England Journal of Medicine* 376 (18): 1704-1706.

Athey, S. and Imbens, G. W. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica* 74 (2), 431-497.

Barrack, R. L., Miller, L. S., Sotile, W. M., Sotile, M. O., and Rubash, H. E. 2006. "Effect of Duty Hour Standards on Burnout among Orthopaedic Surgery Residents." *Clinical Orthopaedics and Related Research* 449: 134-137.

Becker, G. S. 1985. "Human Capital, Effort, and the Sexual Division of Labor." *Journal of Labor Economics* 3 (1): 33-58.

Bolster, L. and Rourke, L. 2015. "The Effect of Restricting Residents' Duty Hours on Patient Safety, Resident Well-Being, and Resident Education: An Updated Systematic Review." *Journal of Graduate Medical Education* 7 (3): 349–363.

Buske, L. 2004. "Younger Physicians Providing Less Direct Patient Care." *Canadian Medical Association Journal* 170 (8): 1217.

Chen, M. K. and Chevalier, J. A. 2012. "Are Women Overinvesting in Education? Evidence from the Medical Profession." *Journal of Human Capital* 6 (2): 124-149.

Clark, K. B. and Summers, L. H. 1982. "Labour Force Participation: Timing and Persistence." *Review of Economic Studies* 49 (5): 825-844.

Clark, A. E. 1999. "Are Wages Habit-Forming? Evidence from Micro Data." *Journal of Economic Behavior & Organization* 39 (2): 179-200.

Coleman, M. T. and Pencavel, J. 1993. "Trends in Market Work Behavior of Women Since 1940." *Industrial and Labor Relations Review* 46 (4): 653-676.

Crossley, T. F., Hurley, J., and Jeon, S. H. 2009. "Physician Labour Supply in Canada: A Cohort Analysis." *Health Economics* 18 (4): 437-456.

Fargen, K. M. and Rosen, C. L. 2013. "Are Duty Hour Regulations Promoting a Culture of Dishonesty Among Resident Physicians?" *Journal of Graduate Medical Education* 5 (4): 553-555.

Feanny, M. A., Scott, B. G., Mattox, K. L., and Hirshberg, A. 2005. "Impact of the 80-Hour Work Week on Resident Emergency Operative Experience." *American Journal of Surgery* 190 (6): 947-949.

Fehr, E. and Goette, L. 2007. "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment." *American Economic Review* 97 (1): 298-317.

Freeman, R. 1980. "The Exit-Voice Tradeoff in the Labor Market: Unionism, Job Tenure, Quits, and Separations." *Quarterly Journal of Economics* 94 (4): 643-673.

Goldin, C. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091-1119.

Gopal, R., Glasheen, J. J., Miyoshi, T. J., and Prochazka, A. V. 2005. "Burnout and Internal Medicine Resident Work-Hour Restrictions." *JAMA Internal Medicine* 165 (22): 2595-2600.

Gronau, R. 1988. "Sex-Related Wage Differentials and Women's Interrupted Careers – the Chicken or the Egg." *Journal of Labor Economics* 6 (3): 277-301.

Heckman, J. J. and Willis, R. J. 1979. "The Distribution of Lifetime Labor Force Participation of Married Women: Reply to Mincer and Ofek." *Journal of Political Economy*, 87 (1): 203-211.

Hutter, M. M., Kellogg, K. C., Ferguson, C. M., Abbott, W. M., and Warshaw, A. L. 2006. "The Impact of the 80-Hour Resident Workweek on Surgical Residents and Attending Surgeons." *Annals of Surgery* 243 (6): 864-875.

Jones, A. M. and Jones, K. B. 2007. "The 88-Hour Family: Effects of the 80-Hour Work Week on Marriage and Childbirth in a Surgical Residency." *Iowa Orthopaedic Journal* 27: 128-133.

Karamanoukian, R. L., Ku, J. K., DeLaRosa, J., Karamanoukian, H. L., and Evans, G. R. 2006. "The Effects of Restricted Work Hours on Clinical Training." *American Surgeon* 72 (1): 19-21.

Kleven, H., Landais, C., and Søgaard, J. E. 2018. "Children and Gender Inequality: Evidence from Denmark." NBER Working Paper 24219.

Kuo, Y.-F., Sharma, G., Freeman, J. L., and Goodwin, J. S. 2009. "Growth in the Care of Older Patients by Hospitalists in the United States." *New England Journal of Medicine* 360 (11): 1102-1112.

Landrigan, C. P., Barger, L. K., Cade, B. E., Ayas, N. T., and Czeisler, C. A. 2006. "Interns' Compliance with Accreditation Council for Graduate Medical Education Work-Hour Limits." *Journal of the American Medical Association* 296 (9): 1063-1070.

Landrigan, C. P., Fahrenkopf, A. M., Lewin, D., Sharek, P. J., Barger, L. K., Eisner, M., Edwards, S., Chiang, V. W., Wiedermann, B. L., and Sectish, T. C. 2008. "Effects of the Accreditation Council for Graduate Medical Education Duty Hour Limits on Sleep, Work Hours, and Safety." *Pediatrics* 122 (2): 250-258.

Martini, S., Arfken, C. L., and Balon, R. 2006. "Comparison of Burnout among Medical Residents Before and After the Implementation of Work Hours Limits." *Academic Psychiatry* 30 (4): 352-355.

McBurney, P. G., Gustafson, K. K., and Darden, P. M. 2008. "Effect of 80-Hour Workweek on Continuity of Care." *Clinical Pediatrics* 47 (8): 803-808.

Melly, B. and Santangelo, G. 2015. "The Changes-in-Changes Model with Covariates." Working paper.

Mincer, J. and Polachek, S. 1974. "Family Investments in Human Capital: Earnings of Women." *Journal of Political Economy* 82 (2): 76-108.

Nakayama, D. K., Thompson, W. M., Wynne, J. L., Dalton, M. L., Bozeman, A. T., and Innes, B. J. 2009. "The Effect of ACGME Duty Hour Restrictions on Operative Continuity of Care." *American Surgeon* 75 (12): 1234-1237.

Philibert, I., Friedmann, P., and Williams, W. T. 2002. "New Requirements for Resident Duty Hours." *Journal of the American Medical Association* 288 (9): 1112-1114.

Philibert, I., Chang, B., Flynn, T., Friedmann, P., Minter, R., Scher, E., and Williams, W. T. 2009. "The 2003 Common Duty Hour Limits: Process, Outcome, and Lessons Learned." *Journal of Graduate Medical Education* 1 (2): 334-337.

Philibert, I., Nasca, T., Brigham, T., and Shapiro, J. 2013. "Duty-Hour limits and Patient Care and Resident Outcomes: Can High-Quality Studies Offer Insight into Complex Relationships?" *Annual Review of Medicine* 64: 467-483.

Polachek, S. 1975. "Potential Biases in Measuring Male-Female Discrimination." *Journal of Human Resources 10* (2): 205-229.

Rizzo, J. and Blumenthal, D. 1994. "Physician Labor Supply: Do Income Effects Matter?" *Journal of Health Economics* 13: 433-453.

Rizzo, J. and Zeckhauser, R. 2007. "Pushing Incomes to Reference Points: Why Do Male Doctors Earn More?" *Journal of Economic Behavior and Organization* 63 (3): 514-536.

Sarma, S., Thind, A., and Chu, M. K. 2011. "Do New Cohorts of Family Physicians Work Less Compared to Their Older Predecessors? The Evidence from Canada." *Social Science & Medicine* 71: 2049-2058.

Sasser, A. 2005. "Gender Differences in Physician Pay: Tradeoffs between Career and Family." *Journal of Human Resources* 40 (2): 477-504.

Scott, A. 2006. "The Productivity of the Health Workforce." *Australian Economic Review* 39 (3): 312-317.

Staiger, D., Auerbach, D., and Buerhaus, P. 2009. "Comparison of Physician Workforce Estimates and Supply Projections." *Journal of the American Medical Association* 302 (15): 1674-1679.

Staiger, D., Auerbach, D., and Buerhaus, P. 2010. "Trends in the Work Hours of Physicians in the United States." *Journal of the American Medical Association* 303 (8): 747-753.

Szymczak, J. E., Brooks, J. V., Volpp, K. G., and Bosk, C. L. 2010. “To Leave or to Lie? Are Concerns about a Shift-Work Mentality and Eroding Professionalism as a Result of Duty-Hour Rules Justified?” *Milbank Quarterly* 88 (3): 350-381.

Vidyarthi, A. R., Arora, V., Schnipper, J. L., Wall, S. D., and Wachter, R. M. 2006. “Managing Discontinuity in Academic Medical Centers: Strategies for a Safe and Effective Resident Sign-Out.” *Journal of Hospital Medicine* 1 (4): 257-266.

Wang, C. and Sweetman, A. 2013. “Gender, Family Status and Physician Labour Supply.” *Social Science & Medicine* 94: 17-25.

Wasserman, M. 2018. “Hours Constraints, Occupational Choice, and Gender: Evidence from Medical Residents.” Working paper.

Watson, D., Slade S., Buske, L., and Tepper, J. 2006. “Intergenerational Differences in Workloads among Primary Care Physicians: A Ten-Year, Populations-Based Study.” *Health Affairs* 25 (6): 1620-1628.

Zhan, C., Miller, M. R., Wong, H., and Meyer, G. S. 2004 “The Effects of HMO Penetration on Preventable Hospitalizations.” *Health Services Research* 39 (2): 345-361.

Figure 1.1: Average Hours Worked per Week by Residents, 1976-2017, Using the ASEC Data

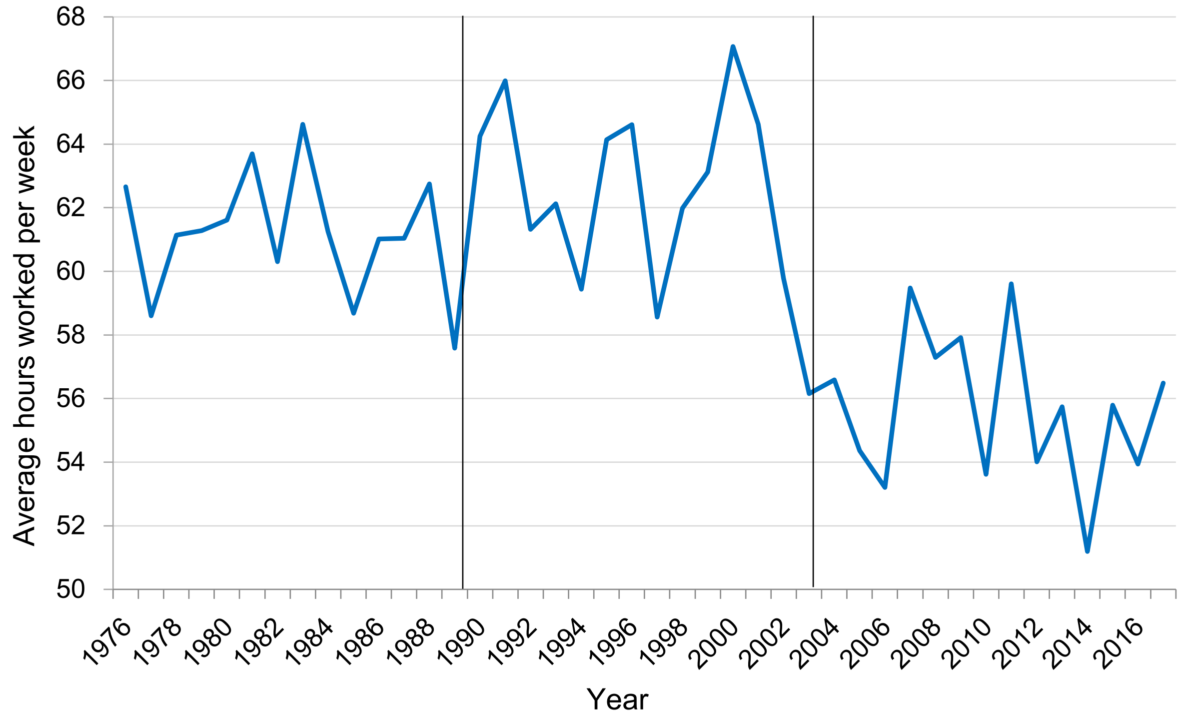


Figure 1.2: Average Hours Worked per Week, 1989-2017

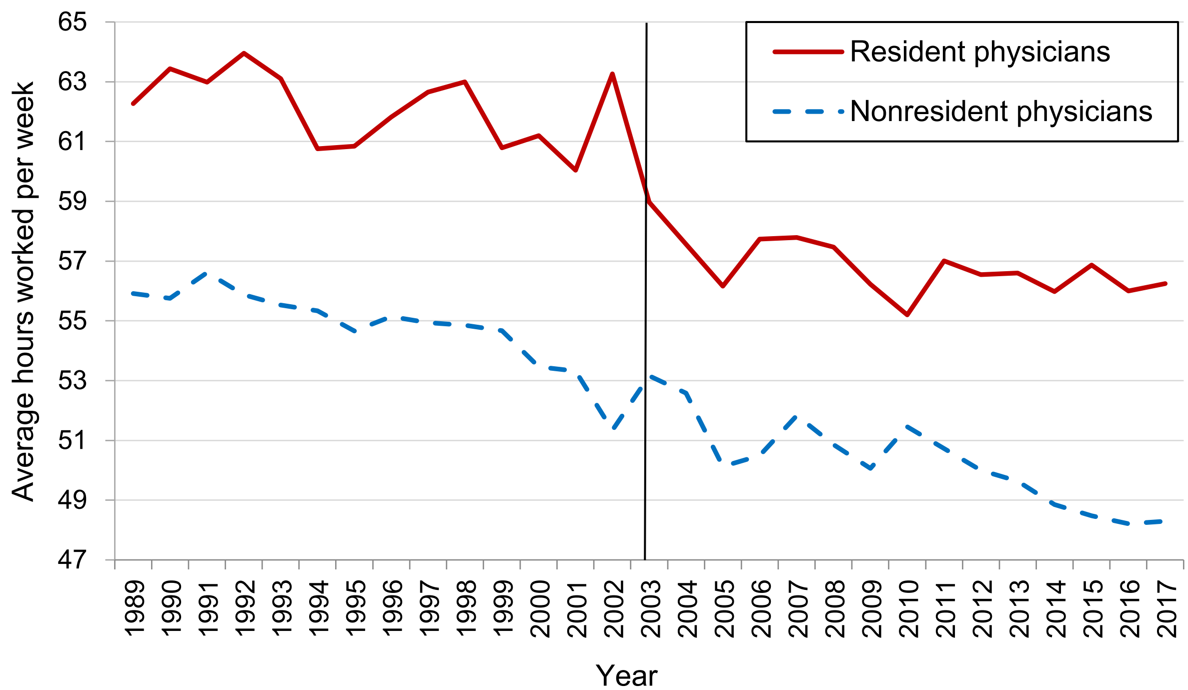


Figure 1.3: Age-Hours Profiles by Treatment Status

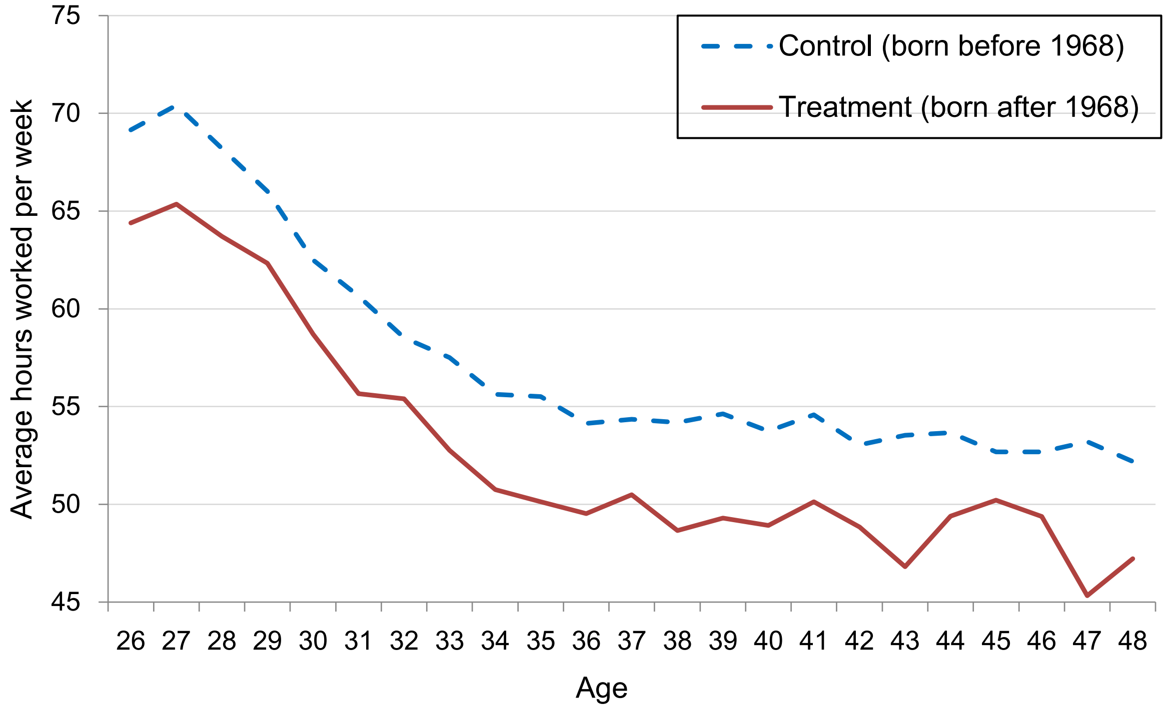


Figure 1.4: Average Hours Worked per Week by Treatment Status

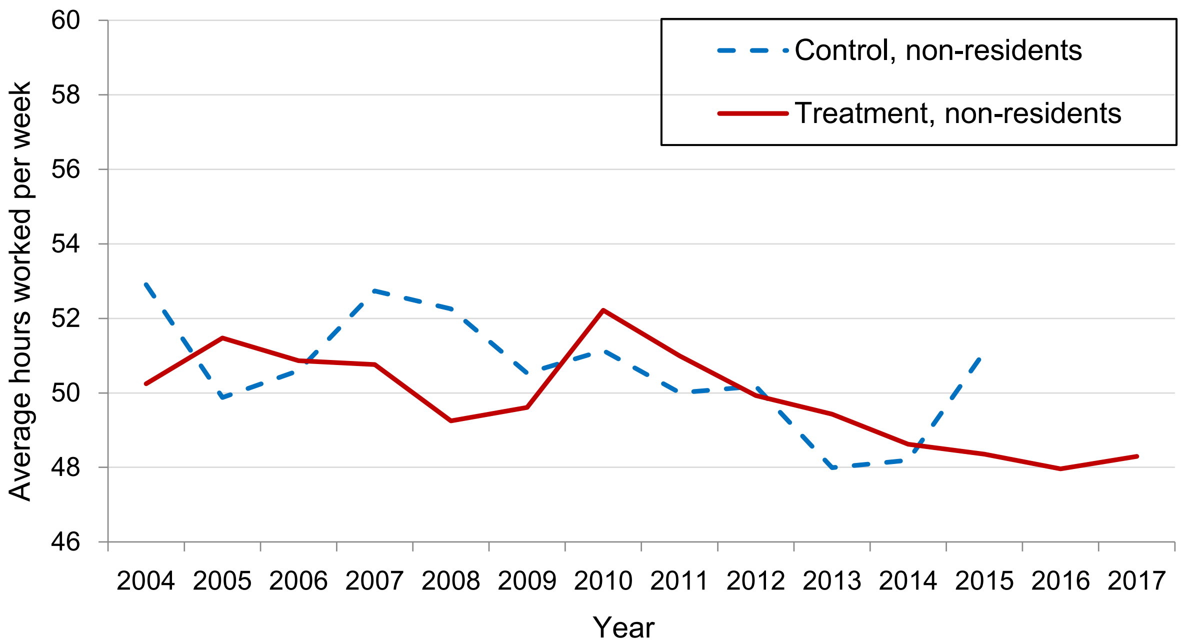
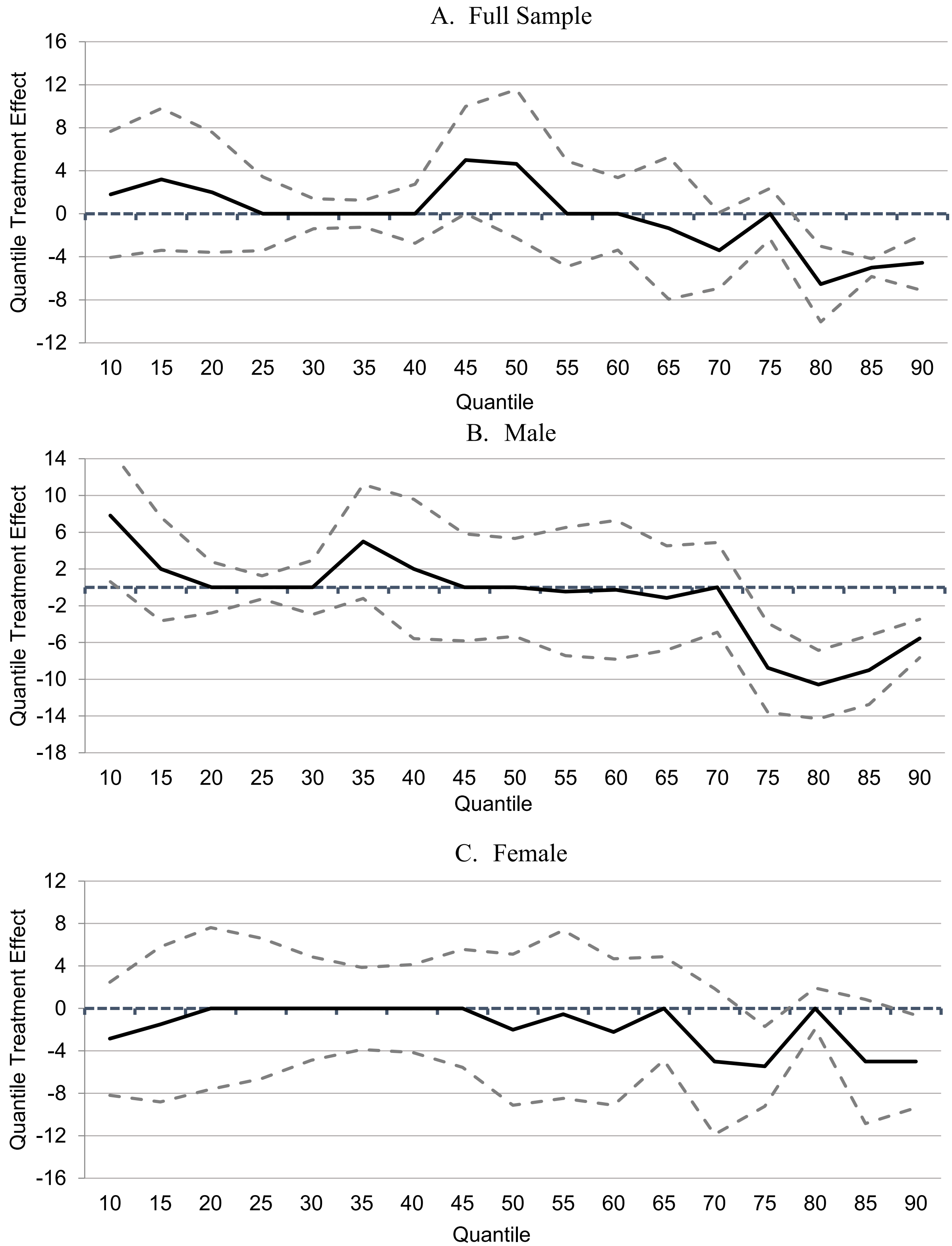


Figure 1.5: Changes-in-Changes Estimates



Notes: These figures show the changes-in-changes estimates of the reform on long-term labor supply for seventeen quantile values, controlling for age, gender, and race/ethnicity. Dotted lines provide bootstrapped pointwise 95-percent confidence intervals for quantile treatment effects with 1,000 replications.

Table 1.1: Summary Statistics

	All	Gender		Treatment Status	
		Male	Female	Treatment	Control
Age	37.608 (6.228)	38.090 (6.209)	36.672 (6.159)	34.192 (5.262)	40.219 (5.606)
Male	0.660 (0.474)			0.575 (0.494)	0.731 (0.444)
White	0.697 (0.459)	0.721 (0.448)	0.651 (0.477)	0.623 (0.485)	0.757 (0.429)
Black	0.059 (0.235)	0.047 (0.211)	0.083 (0.275)	0.068 (0.253)	0.050 (0.219)
Hispanic	0.057 (0.231)	0.057 (0.233)	0.055 (0.229)	0.063 (0.244)	0.054 (0.226)
Asian	0.181 (0.385)	0.169 (0.375)	0.204 (0.403)	0.239 (0.426)	0.134 (0.340)
Other race	0.006 (0.077)	0.005 (0.073)	0.007 (0.083)	0.006 (0.079)	0.006 (0.074)
Married	0.757 (0.429)	0.783 (0.412)	0.706 (0.456)	0.694 (0.461)	0.806 (0.395)
Number of children	1.318 (1.290)	1.439 (1.335)	1.083 (1.163)	0.980 (1.183)	1.579 (1.312)
Family Size	3.193 (1.565)	3.335 (1.599)	2.919 (1.459)	2.820 (1.479)	3.477 (1.570)
Weekly hours worked	54.616 (18.373)	56.974 (17.645)	50.033 (18.887)	54.404 (18.527)	54.886 (18.217)
Exposure	0.438 (0.496)	0.380 (0.486)	0.551 (0.497)		
Observations	70,868	46,784	24,084	27,061	41,495

Notes: Table reports means and standard deviations (in parentheses), weighted using sampling weights. The sample includes physicians aged 26-48 in the monthly CPS between 1989 and 2017. The treatment group includes physicians born after 1968, and the control group includes physicians born before 1968.

Table 1.2: The Impact of the Reform on Resident Hours

Outcome: Hours worked at ages 26-34	All	Male	Female
Exposed	-8.384***	-10.029***	-6.868***
	(0.952)	(1.392)	(2.080)
Age	-2.039***	-1.837***	-2.367***
	(0.095)	(0.131)	(0.190)
Male	4.501***		
	(0.541)		
Black	-0.895	-2.695*	0.542
	(0.743)	(1.464)	(0.976)
Hispanic	0.317	0.460	0.480
	(0.910)	(1.062)	(1.311)
Asian	-2.813***	-2.850***	-2.559**
	(0.614)	(0.620)	(1.074)
Other race	4.841*	4.428	4.182
	(2.610)	(4.738)	(3.219)
Cohort FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
State FEs	Yes	Yes	Yes
Observations	22,708	13,615	9,093

Notes: The dependent variable is weekly hours worked at ages 26 to 34. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights. Cluster-robust standard errors by cohort are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively.

Table 1.3: Difference-in-differences Estimates of the Impact on Long-Term Labor Supply

	Ages 26-34 (Residency)	Ages 35-48 (Post-Residency)	Difference (%)
A. Full Sample			
Treatment group	58.420	49.346	-9.074 (-15.532%)
Control group	60.733	53.586	-7.147 (-11.768%)
Difference	-2.313	-4.240	
Difference-in-differences [Standard Error]			-1.927 (-3.764%) [0.872]**
B. Male			
Treatment group	60.089	52.462	-7.627 (-12.693%)
Control group	62.635	56.098	-6.537 (-10.437%)
Difference	-2.546	-3.636	
Difference-in-differences [Standard Error]			-1.089 (-2.256%) [1.046]
C. Female			
Treatment group	56.238	44.927	-11.311 (-20.113%)
Control group	56.235	46.557	-9.678 (-17.210%)
Difference	0.003	-1.630	
Difference-in-differences [Standard Error]			-1.633 (-2.903%) [1.167]

Notes: The sample includes physicians aged 26-48 in the monthly CPS between 1989 and 2017. The treatment group includes physicians born after 1968, and the control group includes physicians born before 1968. Data are weighted using CPS sampling weights, and the standard errors are clustered by cohort. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively.

Table 1.4: Estimation Results of the Impact on Long-Term Labor Supply

Outcome: Hours worked	All		Male		Female	
	(1)	(2)	(3)	(4)	(5)	(6)
Exposed*Post	-4.385***	-4.281***	-4.034***	-3.961***	-4.476***	-4.445***
	(0.916)	(0.901)	(0.877)	(0.868)	(1.275)	(1.274)
Age		-0.631***		-0.574***		-0.741***
		(0.061)		(0.066)		(0.089)
Male		7.086***				
		(0.434)				
Black		0.014		-0.525		0.571
		(0.585)		(1.047)		(0.644)
Hispanic		-0.191		-0.441		0.474
		(0.516)		(0.643)		(1.061)
Asian		-2.270***		-2.627***		-1.698***
		(0.374)		(0.456)		(0.573)
Other race		3.959***		3.907		4.265***
		(1.463)		(2.375)		(1.530)
Cohort FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	No	Yes	No	Yes	No	Yes
Observations	68,556	68,556	45,453	45,453	23,103	23,103

Notes: The sample includes physicians aged 26-48 in the monthly CPS between 1989 and 2017. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights. Cluster-robust standard errors by cohort are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively.

Table 1.5: The Impact of the Reform on Marriage and Divorce

	All		Male		Female	
	Married	Divorced	Married	Divorced	Married	Divorced
	(1)	(2)	(3)	(4)	(5)	(6)
Exposed	0.050***	-0.002	0.028	0.007	0.092***	-0.015*
	(0.017)	(0.006)	(0.018)	(0.008)	(0.022)	(0.008)
Age	0.019***	0.004***	0.019***	0.003***	0.018***	0.005***
	(0.001)	(0.0004)	(0.001)	(0.0004)	(0.002)	(0.001)
Male	0.039***	-0.026***				
	(0.011)	(0.004)				
Black	-0.151***	0.0003	-0.134***	0.004	-0.167***	-0.004
	(0.022)	(0.009)	(0.030)	(0.012)	(0.027)	(0.010)
Hispanic	-0.007	0.010	-0.005	0.018	-0.009	-0.008
	(0.016)	(0.009)	(0.017)	(0.013)	(0.028)	(0.010)
Asian	0.039***	-0.026***	0.021	-0.022***	0.068***	-0.032***
	(0.011)	(0.004)	(0.015)	(0.004)	(0.017)	(0.007)
Other race	-0.071	-0.016	0.011	-0.010	-0.196***	-0.030*
	(0.043)	(0.011)	(0.048)	(0.015)	(0.064)	(0.017)
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	68,556	68,556	45,453	45,453	23,103	23,103

Notes: The sample includes physicians aged 26-48 in the monthly CPS between 1989 and 2017. The dependent variable are indicators of being married or divorced on the timing of the survey. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights. Cluster-robust standard errors by cohort are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively.

Table 1.6: The Impact of the Reform on Fertility Decisions

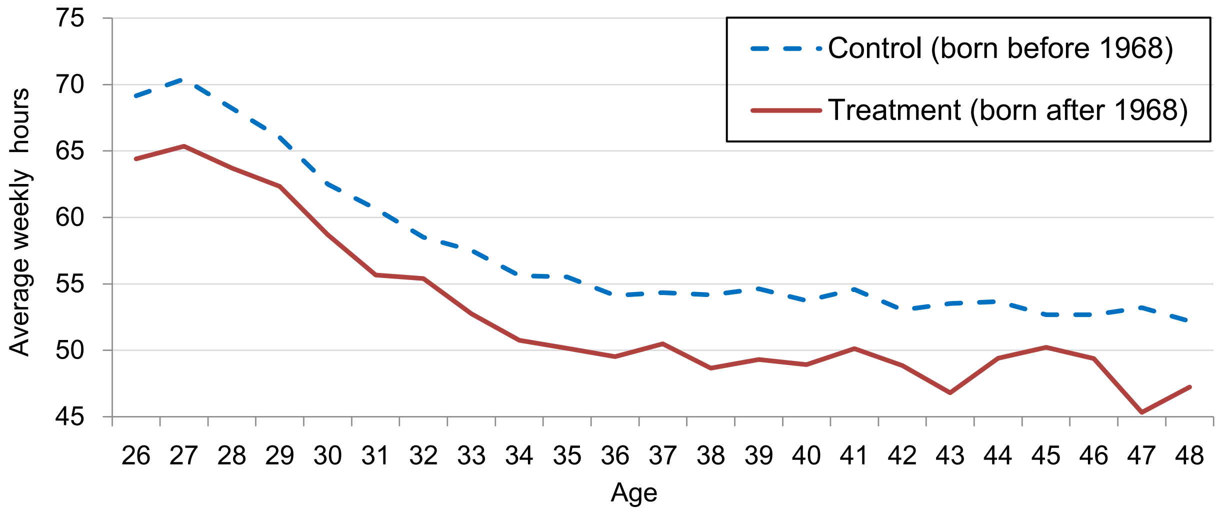
	All		Male		Female	
	Extensive	Intensive	Extensive	Intensive	Extensive	Intensive
	(1)	(2)	(3)	(4)	(5)	(6)
Exposed	0.026 (0.021)	0.039 (0.043)	-0.004 (0.019)	-0.048 (0.047)	0.090** (0.034)	0.202*** (0.059)
Age	0.034*** (0.002)	0.092*** (0.004)	0.032*** (0.002)	0.091*** (0.004)	0.039*** (0.003)	0.092*** (0.005)
Male	0.034*** (0.011)	0.195*** (0.030)				
Black	-0.107*** (0.022)	-0.253*** (0.059)	-0.151*** (0.032)	-0.337*** (0.092)	-0.050** (0.025)	-0.157** (0.065)
Hispanic	-0.032* (0.018)	-0.114** (0.044)	-0.014 (0.019)	-0.060 (0.052)	-0.064** (0.029)	-0.203*** (0.067)
Asian	0.016 (0.011)	-0.078** (0.030)	0.005 (0.014)	-0.118*** (0.041)	0.035* (0.018)	-0.014 (0.040)
Other race	-0.048 (0.047)	-0.173 (0.130)	-0.001 (0.057)	-0.057 (0.185)	-0.117** (0.057)	-0.356** (0.146)
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	68,556	68,556	45,453	45,453	23,103	23,103

Notes: The sample includes physicians aged 26-48 in the monthly CPS between 1989 and 2017. The dependent variables are: (1) fertility at the extensive margin: an indicator of having at least one child and (2) fertility at the intensive margin: the number of children on the time of the survey. White is the omitted group for race/ethnicity variables. Data are weighted using CPS sampling weights. Cluster-robust standard errors by cohort are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively.

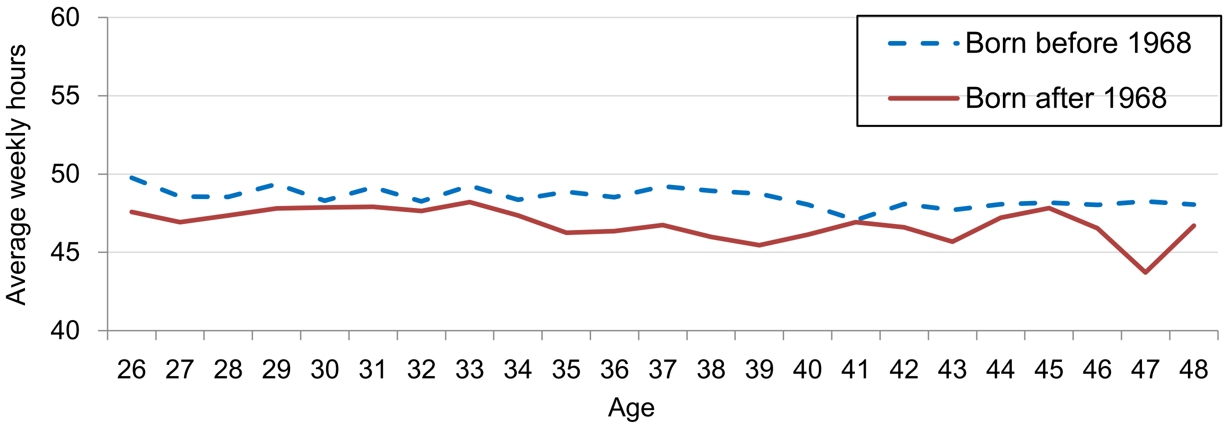
Appendix

Figure 1.A1: Age-Hours Profiles by Cohort

A. Physicians



B. Lawyers, judges, magistrates, and other judicial workers



C. Dentists

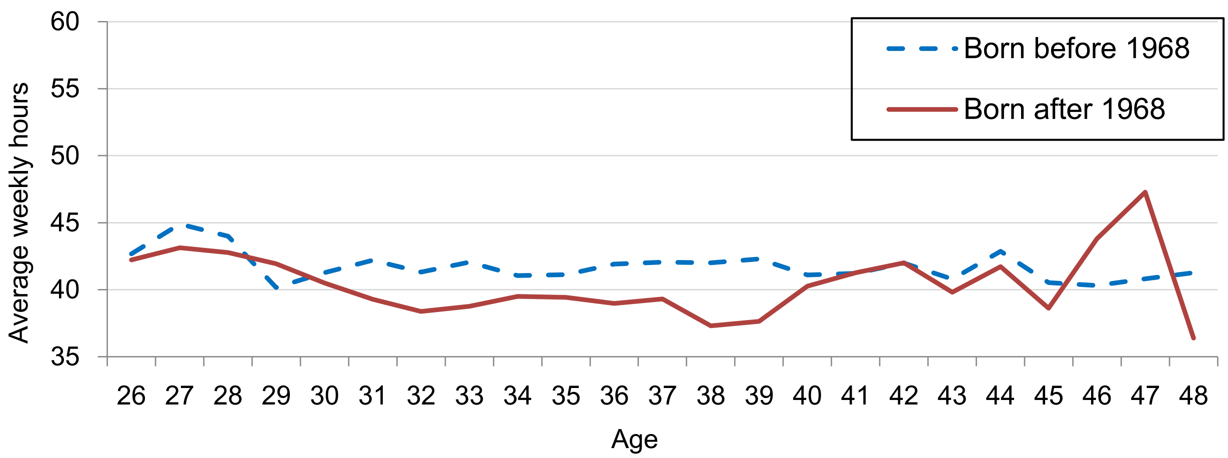


Table 1.A1: Estimation Results of the Impact on Long-Term Labor Supply
– Using Tighter Year-of-Birth Windows for Treatment and Control Groups

Treatment	Control	Eq. (3), no controls	Eq. (3), controls	Observations
All				
After 1968 (Main analysis)	Before 1968	-4.385*** (0.916)	-4.281*** (0.901)	68,556
After 1969	Before 1967	-5.073*** (1.317)	-5.080*** (1.229)	57,856
After 1970	Before 1966	-5.296*** (1.260)	-5.301*** (1.378)	48,087
After 1971	Before 1965	-5.030** (2.154)	-4.850** (2.391)	39,436
Male				
After 1968 (Main analysis)	Before 1968	-4.034*** (0.877)	-3.961*** (0.868)	45,453
After 1969	Before 1967	-4.741*** (1.200)	-4.646*** (1.208)	38,767
After 1970	Before 1966	-5.494*** (1.276)	-5.379*** (1.305)	32,502
After 1971	Before 1965	-4.789*** (1.602)	-4.465*** (1.562)	27,066
Female				
After 1968 (Main analysis)	Before 1968	-4.476*** (1.275)	-4.445*** (1.274)	23,103
After 1969	Before 1967	-5.205*** (1.768)	-5.142*** (1.752)	19,089
After 1970	Before 1966	-4.613** (2.148)	-4.596** (2.118)	15,585
After 1971	Before 1965	-4.957 (4.267)	-4.955 (4.262)	12,370

Notes: Each cell contains an estimate of the effect on post-residency hours. Cluster-robust standard errors by cohort are in parentheses. Data are weighted using CPS sampling weights. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively. Three alternative specifications are used to test the robustness of the results. The treatment status (defined by year of birth) and residency status (defined by age) associated with each row are as follows, where the first one is used in the main analysis.

(1) Treatment: born after 1968; Control: born before 1968. Residency: ages 26-34; Post-residency: ages 35-48.
(2) Treatment: born after 1969; Control: born before 1967. Residency: ages 26-33; Post-residency: ages 36-48.
(3) Treatment: born after 1970; Control: born before 1966. Residency: ages 26-32; Post-residency: ages 37-48.
(4) Treatment: born after 1971; Control: born before 1965. Residency: ages 26-31; Post-residency: ages 38-48.

Table 1.A2: Estimation Results of the Impact on Long-Term Labor Supply
– Excluding Nonrandom Selection into Treatment

	Eq. (3), no controls	Eq. (3), controls	Observations
All			
1941-1991 Cohorts (Main analysis)	-4.385*** (0.916)	-4.281*** (0.901)	68,556
1941-1980 Cohorts	-5.043*** (0.718)	-4.929*** (0.692)	62,819
1941-1978 Cohorts	-5.492*** (0.773)	-5.434*** (0.745)	60,461
1941-1976 Cohorts	-5.948*** (0.835)	-5.909*** (0.803)	57,485
Male			
1941-1991 Cohorts (Main analysis)	-4.034*** (0.877)	-3.961*** (0.868)	45,453
1941-1980 Cohorts	-4.294*** (0.895)	-4.186*** (0.889)	42,471
1941-1978 Cohorts	-4.660*** (0.954)	-4.523*** (0.946)	41,272
1941-1976 Cohorts	-4.677*** (1.029)	-4.575*** (1.020)	39,620
Female			
1941-1991 Cohorts (Main analysis)	-4.476*** (1.275)	-4.445*** (1.274)	23,103
1941-1980 Cohorts	-5.767*** (1.101)	-5.741*** (1.099)	20,348
1941-1978 Cohorts	-6.442*** (1.198)	-6.449*** (1.197)	19,189
1941-1976 Cohorts	-7.556*** (1.292)	-7.556*** (1.289)	17,865

Notes: Each cell contains an estimate of the effect on hours worked after residency. Cluster-robust standard errors by cohort are in parentheses. Data are weighted using CPS sampling weights. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively.

Table 1.A3: Percentiles of the Distribution of Weekly Hours Worked

	Observations	Mean (SD)	Percentiles of the distribution						
			5	10	25	50	75	90	95
All, ages 26-34									
Treatment group	14,919	58.420 (19.078)	32	40	40	60	72	80	90
Control group	7,789	60.733 (20.159)	32	40	47	60	75	90	99
All, ages 35-48									
Treatment group	12,142	49.346 (16.471)	24	32	40	50	60	70	80
Control group	33,706	53.586 (17.493)	25	35	40	50	60	80	84
Male, ages 26-34									
Treatment group	8,215	60.089 (18.566)	36	40	45	60	75	80	95
Control group	5,400	62.635 (19.302)	40	40	50	60	80	94	99
Male, ages 35-48									
Treatment group	7,011	52.462 (16.123)	32	40	40	50	60	75	80
Control group	24,827	56.098 (16.807)	34	40	45	55	65	80	85
Female, ages 26-34									
Treatment group	6,704	56.238 (19.516)	28	36	40	55	70	80	90
Control group	2,389	56.235 (21.398)	23	30	40	52	70	90	99
Female, ages 35-48									
Treatment group	5,131	44.927 (15.940)	20	25	40	40	52	65	75
Control group	8,879	46.557 (17.459)	20	24	40	45	60	70	80

Notes: Table reports means, standard deviations (in parentheses), and percentiles of weekly hours worked, weighted using sampling weights. The sample includes physicians aged 26-48 in the monthly CPS between 1989 and 2017. The treatment group includes physicians born after 1968, and the control group includes physicians born before 1968.

Table 1.A4: The Impact of the Reform on Marriage and Fertility Decisions

		All	Male	Female
Married	Probit	0.050*** (0.015)	0.029* (0.016)	0.089*** (0.021)
	Logit	0.050*** (0.016)	0.030* (0.018)	0.089*** (0.021)
Divorced	Probit	-0.001 (0.006)	0.006 (0.008)	-0.012* (0.007)
	Logit	-0.002 (0.007)	0.005 (0.009)	-0.012* (0.007)
Fertility (Extensive)	Probit	0.026 (0.020)	0.001 (0.017)	0.080** (0.035)
	Logit	0.028 (0.020)	0.005 (0.017)	0.081** (0.036)

Notes: The sample includes physicians aged 26-48 in the monthly CPS between 1989 and 2017. The independent variables include exposure to the regulations, age, gender, race/ethnicity, year fixed effects, and state fixed effects. Data are weighted using CPS sampling weights. Cluster-robust standard errors by cohort are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively. The effects is based on the average marginal effects using probit and logit models.

Table 1.A5: The Impact of the Reform on Marriage and Fertility Decisions

		All	Male	Female
Married	Ages 26-34	0.011 (0.030)	-0.032 (0.045)	0.069 (0.050)
	Ages 35-48	0.044** (0.011)	0.022* (0.012)	0.087*** (0.027)
Divorced	Ages 26-34	0.005 (0.009)	0.015 (0.018)	-0.006 (0.008)
	Ages 35-48	0.0003 (0.006)	0.006 (0.009)	-0.009 (0.011)
Fertility (Extensive)	Ages 26-34	-0.040 (0.034)	-0.074** (0.033)	0.013 (0.049)
	Ages 35-48	0.021 (0.017)	-0.001 (0.018)	0.066* (0.034)
Fertility (Intensive)	Ages 26-34	-0.060 (0.074)	-0.112 (0.074)	0.027 (0.095)
	Ages 35-48	-0.005 (0.057)	-0.081 (0.061)	0.133 (0.083)

Notes: The sample includes physicians aged 26-48 in the monthly CPS between 1989 and 2017. The independent variables include exposure to the regulations, age, gender, race/ethnicity, year fixed effects, and state fixed effects. Data are weighted using CPS sampling weights. Cluster-robust standard errors by cohort are in parentheses. Significance at the 10%, 5%, 1% levels is indicated with 1, 2, 3 asterisks respectively.

Chapter 2

Differential Exposure to Food Insecurity and Government Policy

Alfonso Flores-Lagunes
Department of Economics and Center for Policy Research
Syracuse University, Syracuse, New York, 13244-1020
afloresl@maxwell.syr.edu

Hugo B. Jales
Department of Economics and Center for Policy Research
Syracuse University, Syracuse, New York, 13244-1020
hbjales@maxwell.syr.edu

Judith Liu
Department of Economics and Center for Policy Research
Syracuse University, Syracuse, New York, 13244-1020
jliu61@syr.edu

Norbert L. Wilson
Friedman School of Nutrition Science and Policy
Tufts University, Boston, MA 02111
Norbert.Wilson@tufts.edu

2.1 Introduction

Food insecurity is a serious and persistent problem in the United States, and the prevalence of food insecurity varies considerably among households with different demographic and socioeconomic characteristics (Coleman-Jensen et al., 2018). As shown in Figure 2.1, black- and Hispanic-headed households perennially have substantially higher rates of food insecurity (e.g. 21.8% and 18% in 2017, respectively) than white-headed households (8.8% in 2017).[19] Previous studies have shown that the health outcomes associated with food insecurity are related to children's long-term cognitive and non-cognitive skills that affect human capital investments (e.g., Alaimo et al., 2001; Currie, 2006; Currie, 2009; Almond et al., 2011) and ultimately adult earnings (e.g., Currie, 2009; Ratcliffe, 2015; Bellani and Bia, 2016). Therefore, differential exposure to food insecurity early in life has the potential to heighten and preserve economic inequality. In addition, food insecurity is a likely contributing factor to the disadvantage of those relevant subgroups (e.g., blacks, Hispanics, and immigrants) (e.g., Ratcliffe, 2015; Coleman-Jensen et al., 2016). These groups have received considerable attention from policymakers and academics since some of them have higher rates of poverty and use of public programs at rates greater than the majority populations (e.g., Currie, 2003; Jensen, 2002; Ratcliffe, 2015).

The Supplemental Nutrition Assistance Program (SNAP, formerly known as the Food Stamp Program) is the largest food-assistance program in the U.S. While it is designed to alleviate food insecurity, the program also serves to mitigate the consequences of poverty. Many studies have addressed how SNAP lowers food insecurity overall (e.g., Wilde and Nord, 2005; Gundersen, et al. 2011; Ratcliffe et al., 2011), but little is known about how SNAP mitigates differences in food insecurity exposure by race/ethnicity and other demographic characteristics.

[19] The measure of food insecurity captures household access to food, which may be different from actual nutrition intake.

Our previous study (Flores-Lagunes et al., 2018) finds that SNAP does not have a significant impact on group differences in food insecurity exposure but has potentially different pathways to affect food insecurity for different populations.

This paper aims to provide a closer examination of the role of SNAP on differential exposure to food insecurity. We develop a sequential framework that helps understand how policy rules that are not designed to take into consideration group membership, such as race and ethnicity, can yet have different effects across groups. We decompose the differences in SNAP benefit levels across racial/ethnic groups into three components: (1) the eligibility component – the proportion of households that are eligible for SNAP; (2) the participation component – the propensity to enroll in SNAP among eligible households; (3) the generosity component – the magnitude of SNAP benefits that eligible and participating households receive. Since policy makers are ultimately interested in differences in food consumption and food insecurity exposure, we show that differences in SNAP benefit levels can be linked to differences in food consumption through a factor of proportionality given by the marginal propensity to consume food (MPCF) out of SNAP benefits. The relative importance of differences in eligibility, participation, and generosity obtained by our decomposition remain the same regardless of whether we are looking at SNAP benefit levels or food consumption outcomes.

We use data from the December Current Population Survey (CPS) between 2003 and 2016, along with its Food Security Supplement (FSS), and impute SNAP eligibility and benefits for the analysis sample. We decompose group differences in SNAP benefits into the three components described above. Our main findings are as follows. First, we show that differences in the proportion of being eligible alone can explain a substantial part of the total difference in the mean benefits for both black-white and Hispanic-white differentials. Second, participation leads to an upward shift of the benefit levels for blacks, but it has little impact for Hispanics. In contrast, the generosity component increases Hispanics' SNAP benefits, but it is negligible for blacks. Third, combining our results with the estimated MPCF out of SNAP benefits in Hastings and Shapiro (2018), SNAP reduces the food consumption gaps between blacks/Hispanics and whites by a modest amount, which is likely not enough to reduce the differences in the resource gaps between groups.

We also provide an exploratory analysis of how changes to SNAP policy rules might affect differences in food insecurity across groups by examining three counterfactual policy scenarios. The first scenario is constant transfer, which provides all participants the same amount of SNAP benefits. The second scenario involves automatic enrollment, which makes all eligible households automatically enroll in the SNAP program. The third scenario is universal eligibility, which makes all households become eligible for SNAP. Among these policy counterfactuals, automatic enrollment raises both blacks' and Hispanics' benefit levels relative to whites the most. The constant transfer policy slightly increases benefits for blacks relative to whites, but it substantially lowers the differences in benefit levels between Hispanics and whites, compared to the baseline decomposition. Universal eligibility has little impact on the differences between blacks and whites, as well as Hispanics and whites. Overall, our results suggest that, among the three exploratory policy scenarios, the automatic enrollment policy may be the most effective in alleviating differences in exposure to food insecurity across racial and ethnic groups.

Overall, this paper contributes to our broader understanding of policies designed to reduce poverty in several ways. First, we uncover the pathways through which SNAP may have on the existing heterogeneity in exposure to food insecurity over the demographic groups under consideration. Second, we provide a framework that conceptualizes how a color-blind program, like SNAP, could have differences in benefit levels across groups and affect inequality. The structural model developed in this paper allows us to parse out different components of the program and tell what would happen to inequality when a policy change occurs. Third, this technique can be applied to other social programs and can help policy makers target policies more effectively to alleviate social and economic inequality.[20]

The rest of the paper is structured as follows. Section 2 provides background on the Supplemental Nutrition Assistance Program (SNAP). Section 3 describes the model used to decompose differences in SNAP benefit levels and shows how it can be linked to differences in food consumption. Section 4 describes the data. Section 5 presents descriptive statistics and results. Section 6 concludes and discusses future work.

2.2 Background

In 2017, about 15 million households were food insecure in the U.S., including 5.8 million with "very low" food security (Coleman-Jensen et al., 2018). The main policy lever against exposure to food-related hardship in the U.S. is the Supplemental Nutrition Assistance Program (SNAP, formerly known as the Food Stamp Program). In fiscal year 2017, SNAP provided benefits to 20.8 million households at a cost of $68 billion, which is much larger than any other food and nutrition assistance programs such as Women Infants and Children (WIC), the National School Lunch Program (NSLP), and the School Breakfast Program (USDA, 2019). In addition, SNAP is the most universal and unrestricted food-assistance program: virtually available to all households that meet the financial and nonfinancial eligibility criteria.

[20] For instance, our decomposition analysis can be used to investigate how states' economic and policy environments shape the educational disparities in mortality rates over time, as found in Montez et al. (forthcoming).

To be eligible for SNAP, a typical household must meet three financial criteria, while households with a disabled member or a member whose age is 60 or above face less stringent criteria. The criteria are: (1) gross monthly income does not exceed l30 percent of the poverty line (or 165 percent of the poverty line for households with an elderly or disabled member); (2) net monthly income is at or below the poverty line, where net income is calculated as gross income minus allowable deductions (including a 20 percent deduction from gross income, a standard deduction, a deduction for households incurring expenses in the care of their children and/or disabled dependents, a medical deduction for expenses, and a shelter deduction for costs above 50% of a household's net income, computed before the shelter deduction and capped except for elderly or disabled households); (3) countable assets are no more than $2,250 (or $3,250 for households with an elderly or disabled member). Besides these three main criteria, able-bodied adults without dependents are limited to receiving benefits for 3 months out of each 36-month period if they do not meet certain work requirements.

SNAP provides monthly benefits to eligible households to purchase food items at SNAP-authorized retailers with an electronic benefit transfer (EBT) card.[21] The monthly SNAP benefit amount is the maximum SNAP allotment, varied by household size, less 30 percent of a household's net monthly income, and the benefit amount is subject to a minimum amount.[22] In fiscal year 2017, the average benefit level was about $254 per household per month (USDA, 2019). To receive SNAP benefits, program applicants must provide required documentation and

[21] Since June of 2004, all States have implemented the Electronic Benefits Transfer (EBT) system.
[22] Note that the poverty line is nonlinearly related to household size and composition. The USDA adjusts the income eligibility standards, the deductions, and the maximum allotments at the beginning of each fiscal year, which takes effect from October 1st of the previous year to September 30th of the current year. The changes are based on changes in the cost of living. These parameters are the same for all states in the continental U.S. but different for Alaska and Hawaii.

participate in an interview. After initial eligibility, recipients must be recertified every 6 to 24 months.

Coleman-Jensen et al. (2018) points out that the rates of food insecurity are higher than the national average for certain populations.[23] In particular, blacks and Hispanics have substantially higher rates of food insecurity than whites over time (Figure 2.1).[24] Using data from Oklahoma, Nam et al. (2015) also shows that whites experience significantly lower incidence of food insecurity than the minority groups (African Americans, American Indians, and Hispanics). To better understand the differential exposure to food insecurity, our previous study (Flores-Lagunes et al., 2018) uses the Oaxaca-Blinder decomposition to assess the contribution of different factors to the observed group differences in food insecurity incidence and severity by race, ethnicity, and immigrant status. These factors are an "endowment component," attributable to group differences in observable household characteristics, and a "structural component," attributable to group differences in the structure linking the observable household characteristics to food insecurity. Our finding is suggestive of the heterogeneity in the relative importance of the factors (endowment and structure) contributing to the observed differences in food insecurity exposure across these demographic groups.

How does SNAP affect the inequality in food insecurity exposure? Unlike other food and nutrition programs (e.g., WIC and NSLP), SNAP is universal in the sense that it is not targeted at specific demographic groups; i.e., it is blind to race and ethnicity. However, different racial/ethnic groups might have different household characteristics and participate in SNAP at

[23] For example, black- and Hispanic-headed households, low-income households, single-headed households with children.

[24] Note that this is based on all households in the U.S. Similarly, Flores-Lagunes et al. (2018) shows that blacks and Hispanics have substantially higher rates of food insecurity than whites using the target population in our analysis, which will be introduced in Section 2.4.

different rates, which results in different SNAP receipt and benefit levels. For instance, one possible explanation to the disparate patterns in food insecurity among the groups analyzed is that they may participate in SNAP at different rates. Therefore, SNAP could potentially have different pathways to affect food insecurity for different populations. Motivated by this, we further take a closer examination of the role played by SNAP and the potentially different determinants of program participation across groups.

2.3 Decomposition of SNAP

To understand the role of SNAP on differential exposure to food insecurity over the demographic groups under consideration, we decompose group differences in SNAP benefit levels into three components: (1) the eligibility component – the proportion of households that are eligible for SNAP; (2) the participation component – the propensity to enroll in SNAP among eligible households; (3) the generosity component – the magnitude of SNAP benefits that eligible and participating households receive.

2.3.1 Baseline model

We consider a population with two non-overlapping subgroups indexed by $g \in \{0,1\}$. Let 0 denotes the population of whites, and 1 denotes the population of blacks. To fix ideas, we start with the simplest case of the decomposition analysis, with binary indicators of SNAP eligibility and generosity. For any household in group g, we observe a binary variable for eligibility, L_g, where $L_g = 1$ if a household is eligible for SNAP, and $L_g = 0$ if a household is not eligible for SNAP. Among eligible households, we observe whether eligible households take up their SNAP benefits or not, captured by a dummy variable T_g. It equals 1 if a household is enrolled in SNAP, and 0 otherwise. For those who participate in SNAP, some of them qualify for a high amount of SNAP benefits, Y, while others qualify for a low amount of benefits which, for the purposes of this example, we set it to $\frac{Y}{2}$. Let H_g be an indicator of receiving the high amount, which equals 1 if a participating household receives the high amount, and 0 otherwise. Therefore, the expected value of SNAP benefits for group g can be calculated as:

$$E\left[Y_g\right] = Pr\left(L_g = 1\right) \cdot Pr\left(T_g = 1 | L_g = 1\right) \cdot [Pr\left(H_g = 0 | L_g = 1, T_g = 1\right) \cdot \frac{Y}{2} +$$

$$Pr\left(H_g = 1 | L_g = 1, T_g = 1\right) \cdot Y].$$

The difference in the mean SNAP benefit levels between groups is

$$\Delta_Y = E[Y | g = 1] - E[Y | g = 0].$$

We can decompose this overall difference into the following three components related to SNAP:

(1) Eligibility: the group difference in the proportions of households that are eligible. To obtain this, we calculate the counterfactual expected outcome where group 1 has the same eligibility rate as group 0:

$$E[Y^{C1} | g = 1] = Pr(L_0 = 1) \cdot Pr(T_1 = 1 | L_1 = 1) \cdot \Big[Pr(H_1 = 0 | L_1 = 1, T_1 = 1) \cdot \frac{Y}{2} +$$

$$Pr(H_1 = 1 | L_1 = 1, T_1 = 1) \cdot Y\Big].$$

The contribution of eligibility to the overall difference is equal to

$$\Delta_L = E[Y | g = 1] - E[Y^{C1} | g = 1],$$

where the two terms on the right-hand side are the same except for the part corresponding to the eligibility component. As is clear from the above, the contribution of this component will disappear if group 1 and group 0 have the same proportion of households being eligible for SNAP.

(2) Participation: he group difference in the participation rates. To obtain this, we calculate the counterfactual expected outcome where group 1 has the same eligibility and take-up rates as group 0:

$$E[Y^{C2}|g=1] = Pr(L_0=1) \cdot Pr(T_0=1|L_0=1) \cdot \Big[Pr(H_1=0|L_1=1,T_1=1) \cdot \frac{Y}{2} + Pr(H_1=1|L_1=1,T_1=1) \cdot Y\Big].$$

The contribution of participation to the overall difference (Δ_T) is equal to

$$E[Y^{C1}|g=1] - E[Y^{C2}|g=1],$$

where everything is equal except for the part corresponding to participation component.

(3) Generosity: the group difference in the proportions of households who qualify for the high amount. To obtain this, we calculate the counterfactual expected outcome where group 1 has the same amount of benefits as group 0: $E[Y|g=0]$, and the contribution of generosity to the overall difference (Δ_G) is equal to

$$E[Y^{C2}|g=1] - E[Y|g=0].$$

Δ_G will become zero if group 1 and group 0 receive the same amount of SNAP benefits. Altogether, the three group differences add up to the overall difference, that is, $\Delta_Y = \Delta_L + \Delta_T + \Delta_G$. Since the eligibility and generosity components are based on predetermined and known policy rules of the SNAP program, group differences in these two components are deterministic and entirely due to differences in the composition of these groups.

2.3.2 Numerical Example

We use the following numerical example to illustrate how this decomposition works. Suppose the high amount of SNAP benefits (Y) is equal to 500 dollars per month. For households in group 1, 50 percent are eligible for SNAP, among which 50 percent enroll and a fraction of 0.3 qualify for the high amount of benefits. For households in group 0, 30 percent are eligible for SNAP, among which 60 percent enroll and one-fourth qualify for the high amount of benefits. The above information is summarized in Table 2.1.

Hence, the expected SNAP benefits for group 1 is equal to

$$E[Y|g = 1] = 0.5 \times 0.5 \times (0.7 \times 250 + 0.3 \times 500) = 81.25,$$

and the expected SNAP benefits for group 0 is equal to

$$E[Y|g = 0] = 0.3 \times 0.6 \times (0.75 \times 250 + 0.25 \times 500) = 56.25.$$

The overall difference in the average SNAP benefits is $\Delta_Y = E[Y|g = 1] - E[Y|g = 0] = 25$, which means that blacks receive 25 dollars more of SNAP benefits per month than whites on average.

Using the counterfactual expected outcomes, we can calculate the contributions of eligibility, participation, and generosity to the overall difference. First, the contribution of eligibility is equal to

$$\Delta_L = 81.25 - 0.3 \times 0.5 \times (0.7 \times 250 + 0.3 \times 500) = 32.5.$$

Second, the contribution of participation is equal to

$$\Delta_T = 0.3 \times 0.5 \times (0.7 \times 250 + 0.3 \times 500) - 0.3 \times 0.6 \times (0.7 \times 250 + 0.3 \times 500) = -9.75.$$

Third, the contribution of generosity is equal to

$$\Delta_G = 0.3 \times 0.6 \times (0.7 \times 250 + 0.3 \times 500) - 56.25 = 2.25.$$

Based on this example, the decomposition of the overall difference in SNAP benefit levels between groups shows that the three components contribute to the black-white differential differently, with the eligibility component being somewhat more important than the other two. The participation component contributes a negative effect to the overall difference. Altogether, these three components sum up to the overall difference:

$$\Delta_L + \Delta_T + \Delta_G = 32.5 - 9.75 + 2.25 = 25.$$

In addition to the decomposition, this model also allows us to evaluate counterfactual policy scenarios. We consider the following three scenarios: (1) constant transfer: assuming all the SNAP participants receive the high amount of 500 dollars per month; i.e., $Pr\left(H_g = 1 | L_g = 1, T_g = 1\right) = 1$; (2) automatic enrollment: assuming all eligible households enroll in SNAP; i.e., $Pr\left(T_g = 1 | L_g = 1\right) = 1$; (3) universal eligibility: assuming all households are eligible for SNAP; i.e., $Pr\left(L_g = 1\right) = 1$. Table 2.2 shows the calculated effects of changes in SNAP policy rules on group differences in SNAP benefits and the three components, in comparison with the baseline result shown in Column (1). As seen in Columns (2)-(4), automatic enrollment increases the average benefit level of blacks relative to whites the most, and this is mostly attributable to the participation component. Constant transfer also raises the average benefit level of blacks, in which the eligibility and participation components have contradicting effects, with the former outweighs the latter. In contrast, universal eligibility decreases SNAP benefits of blacks relative to whites, and this is mainly attributable to the participation component.

2.3.3 General Setting

In this section, we discuss more formally and generally the conditions required for this decomposition to work. Fortin et al. (2011) show that standard decomposition methods, such as Oaxaca and Blinder, rely on the assumption of ignorability. Ignorability essentially imposes that, conditional on the value of observable characteristics, the distribution of potential outcomes is independent of the treatment status. Here, we impose a modified version of this ignorability assumption.

Let $Y(1)$ be the potential value of the SNAP benefit a household receive if the household belongs to group 1, and $Y(0)$ be the counterfactual value of the SNAP benefit that the household could have receive if the household were to belong to group 0. Let X be the variables that unequivocally determine both eligibility and generosity of the program. Our key assumption is then:

$$Y(1) = Y(0)|T, L, X.$$

This assumption essentially restricts eligibility and benefit levels to be invariant to group membership. That is, conditional on eligibility and participation, SNAP benefits are not different between groups. This assumption can be easily verified by a casual look at the program rules. Imposing this condition allows us to bypass the need to identify the structural component of standard decompositions since this component is trivially equal to zero. That is, all of the differences in $E[Y|G = g]$ between groups must be accounted by composition effects.

As argued by Fortin et al. (2011), the structural component of a decomposition exercise can be thought as a treatment effect (of group membership on the outcome), and the composition effect reflects differences in the distribution of predetermined characteristics between groups. Imposing the assumption above, we immediately obtain the result that the structural component, the treatment effect of group membership on SNAP benefit levels, is zero. Thus, if any differences in average benefit levels are observed between groups, this must be entirely accounted by their differences in the distribution of (T, L, X). It is useful to write the joint distribution of (T, L, X) using the factorization formula:

$$f(t, l, x) = f(x)\Pr[L = l|X = x]\Pr[T = t|L = l, X = x].$$

It is important to note that if X includes all characteristics that affect eligibility, then $\Pr[L = l|X = x]$ becomes degenerate; that is, it must be either zero or one. In addition, we know that T

is zero whenever L is zero and that Y is zero unless T and L are one. Since the deterministic function of (T, L, X), $Y(T, L, X)$, is degenerate, we obtain:

$$Y(T, L, X) = T\big(L(X)\big)L(X)m(X),$$

given a known function $m(X)$. Therefore, the expected value of the SNAP benefit for households in group g is

$$E[Y|G = g] = \Pr[L = 1|G = g] \Pr[T = 1|L = 1, G = g] \, E[Y|L = 1, T = 1, G = g],$$

where $\Pr[L = 1|G = g] = \Pr[X \in A|G = g]$, where A is the set of values of X that determine the eligibility of a household. It is relevant to note also that the set A is not indexed by g, that is, eligibility thresholds are the same regardless of group membership. For the third term in the equation above, we have that

$$E[Y|L = 1, T = 1, G = g] = \int Y(1,1, x)f(x|L = 1, T = 1, G = g)dx,$$

which can be re-written as $\int f(x|L = 1, T = 1, G = g)m(x)dx$.

To sum up, the overall difference in the observed SNAP benefit levels between groups can be decomposed into three components:

$$\Delta_Y = \Delta_L + \Delta_T + \Delta_G,$$

where

$\Delta_L = (\Pr[L = 1|G = 1] - \Pr[L = 1|G = 0]) \Pr[T = 1|L = 1, G = 1] \, E[Y|L = 1, T = 1, G = 1];$

$\Delta_T = \Pr[L = 1|G = 0](\Pr[T = 1|L = 1, G = 1] - \Pr[T = 1|L = 1, G = 0]) \, E[Y|L = 1, T = 1, G = 1];$

$\Delta_G = \Pr[L = 1|G = 0] \Pr[T = 1|L = 1, G = 0] \, (E[Y|L = 1, T = 1, G = 1] - E[Y|L = 1, T = 1, G = 0]).$

Each of these differences captures the mean effect of a counterfactual experiment conducted in group 1 that changes the respective distribution of the component to their corresponding counterpart in group 0 while holding everything else constant. For example, Δ_L answers the counterfactual question of how average SNAP benefit levels would change if group 1 were to have the same eligibility rates as group 0 while maintaining fixed the likelihood of participating in the program and also the average level of transfers that they are entitled to. Similarly, Δ_T answers the counterfactual question of how average SNAP benefit levels would change if, on top of having the same eligibility rates as group 0, we also were to apply the same likelihood of participating in the program as households in group 0. Finally, the term Δ_G answers the question of how average SNAP benefit levels would change if, on top of having the same eligibility rates and participation rates as group 0, we were to entitle households in group 1 to the same levels of average benefits that households in group 0 are entitled to. These three components, differences in eligibility rates, participation rates, and average benefit levels by construction add to the overall difference between groups. In this sense, this decomposition exercise is completely atheoretical.

2.3.4 Linkage to Differential Exposure to Food Insecurity

So far, we undertake a close to mechanical exercise to understand the average difference in SNAP benefit levels between groups. The limitation of this approach is that the outcome of interest is a deterministic function of household characteristics. Moreover, policy makers are ultimately interested in differences in food consumption and food insecurity exposure. Differences in SNAP benefits across groups are only of interest to the extent that these differences can trace out the differences in food consumption and food insecurity exposure. In this section, we argue that the decomposition exercise above can help inform policy makers about the outcomes that are of interest, such as food consumption, even if the object of the decomposition is limited to SNAP benefits.

Assume that food consumption (C) is related to SNAP benefits and other household characteristics according to the following equation:

$$C_{ig} = \beta_{ig} Y_{ig} + \theta X_{ig} + \varepsilon_{ig}.$$

This equation states that food consumption is a function of household characteristics (X_{ig}), the level of SNAP benefits that a household receives (Y_{ig}), and their propensity to consume out of SNAP benefits (β_{ig}), which we allow to vary across households and groups.

Under the assumption that the error term is mean independent of the observed characteristics in the regression, taking the differences of average food consumption levels across groups and adding and subtracting a couple of terms, we obtain:

$$\Delta E[C] = \beta_1 \Delta E[Y] + \Delta\beta E[Y_0] + \theta\Delta X + \Delta Cov(\beta, \mathrm{Y}).$$

This equation looks almost identical to the standard Oaxaca-Blinder decomposition, except for the last term. The last term accounts for the potential differences in the covariance between the propensity to consume out of SNAP benefits and the benefit levels themselves between groups. This term can be safely ignored if (1) $Cov(\beta_i, Y_i)$ is zero for both groups, which must happen if SNAP benefits are constant, (2) propensity to consume out of SNAP is constant among members of the same group, (3) these covariances are the same regardless of group membership, or (4) β_i and Y_i are independent.[25]

Under the assumption that the last term is zero, group differences in SNAP benefit levels that we decompose before will affect differences in food consumption through a factor of

[25] It is possible that β_i is decreasing with Y_i; i.e., MPCF out of SNAP is low when the benefit level is high. We can still take care of this case by estimating the covariance between β_i and Y_i.

proportionality given by the marginal propensity to consume food (MPCF) out of SNAP benefits, which is a parameter that has been the object of study in the previous literature. In particular, the relative importance of differences in eligibility, take-up, and generosity obtained by our decomposition must remain the same regardless of whether we are looking at SNAP benefit levels or food consumption outcomes.

Previous studies have examined what proportion of SNAP receipts are infra-marginal. For instance, Hoynes and Schanzenbach (2009) find that the MPCF out of food stamps is 0.16 for all non-elderly and 0.30 for female-headed households; Bruich (2014) suggests that the MPCF out of food stamps is 0.3; Hastings and Shapiro (2018) use administrative data and casual inference approaches and show that the MPCF out of SNAP benefits is 0.5 to 0.6, which is larger than the MPCF out of cash. In Section 5, we use the result from Hastings and Shapiro (2018), as it is the most recent and reliable study, to link our decomposition results of differences in SNAP benefit levels to differences in food consumption.[26]

2.4 Data

We analyze data from the December Current Population Survey (CPS) between 2003 and 2016, along with the Food Security Supplement (FSS). These data are nationally representative of the U.S. population and include sufficient information on household characteristics that allow us to conduct the decomposition analysis. The unit of observation for the analysis is at the household level. Households are included in the sample if the reference person is above 15 years old. We focus on households with incomes below 185 percent of the poverty line or that report being short of money for food. These are the target population being asked about food insecurity

[26] It is worth noting that different groups might have different levels of MPCF out of SNAP benefits. We assume that they are the same across groups in this paper.

questions and SNAP participation in the FSS. It is important to note that his target population is more economically disadvantaged and more food-insecure than the general population.

Since SNAP eligibility information is not available in the CPS, we impute program eligibility for each household using the data on household income and family composition. To obtain adequate information on earnings, we read in the Outgoing Rotation Group (ORG) files in January-March each year and match the December data to the appropriate ORG.[27] The match may fail because of identifier errors (due to migration, mortality, non-response, and recording errors), inconsistencies in respondents' basic demographic attributes (race, age, or gender), or incomplete information on the key variables. Overall, a total of 160,065 respondents within the scope of our study have complete information on earnings and family income after matching the FSS to the ORG.

We collect the eligibility standards every year from the USDA to impute SNAP eligibility for our sample with the following five steps. First, we use information on weekly earnings to pass through the gross and net monthly income tests. Second, the categorical family income variable is used to further screen out certain ineligible households. Third, we employ different income eligibility standards for disabled and elderly (age 60 or older) respondents. Fourth, we rule out immigrants who have lived in the U.S. for less than five years as they are ineligible. Fifth, households are eligible if they reported participating in SNAP.[28] Note that the December CPS-FSS does not track all of the information needed to identify eligible households. For instance, we lack information on households' assets, expenses related to medical and shelter

[27] For respondents in the December CPS, the ORG is split into December-March CPS surveys. We use CPS identifiers to match households across survey months of January-March.

[28] Among our imputed ineligible households, 6.68 percent (10,698 out of 160,065) of them are shown participating in SNAP, which could be due to misreporting or the lack of information to identify eligibility. We thus consider these households as eligible.

deductions, SSI and TANF receipt, and whether there is a disabled or elderly member in the household other than the respondent. We assume that all types of income other than self-reported earnings and family income are zero.

Our imputed eligibility seems fairly reasonable. Compared to those documented in the USDA reports (e.g., Wolkwitz, 2008; Cunnyngham, 2018), the take-up rates (see Figure 2.2) have similar trends across demographic groups and over time periods. The program shows a countercyclical pattern, increasing in take-up rates notably in the Great Recession. The reduction in take-up rates after 2013 is consistent with the fact that, in November 2013, all SNAP benefits were reduced when temporary increases in the American Recovery and Reinvestment Act expired. In addition, the take-up rates are much lower for eligible elderly adults (age 60 or older) than their counterparts. However, the take-up rates vary substantially across studies, which is mainly due to different data, methodology, and analysis samples used.[29]

Since information on SNAP benefit receipt is available in the December CPS, we can further impute the amount of benefits for those who participate in SNAP using the annual benefit standards from the USDA. Our analysis is based on the imputed benefit levels since there are several major problems with the self-reported SNAP benefits in the CPS. First, the self-reported values seem to have a rounding problem. There are clear spikes in the density at benefit amounts divisible by 100. Second, the SNAP benefits are top-coded in the CPS (the top code is $450 before 2011 and $700 in 2011 and after). Third, there are a number of participants refused, didn’t know, or didn’t response their SNAP benefit amounts.

[29] Our estimated participation rates lie between those in Cunnyngham (2018) and Gundersen et al. (2018). The USDA reports collect administrative data from the SNAP Quality Control data to get information on SNAP participation, along with the data from the CPS Annual Social and Economic Supplement to generate SNAP eligibility (e.g., Cunnyngham, 2018). Compared with their estimates, our take-up rates seem to be low, but the pattern is pretty similar. Previous studies have pointed out that SNAP participation is somewhat underreported in survey data (Gundersen and Kreider, 2008; Cunnyngham, 2018). In addition, the USDA sample is different from our sample.

2.5 Results

2.5.1 Descriptive Statistics

Table 2.3 describes the summary statistics for the analysis sample by race and ethnicity. As shown in Panel A, whites have substantially higher monthly earnings and family income than blacks and Hispanics on average. Using whites as the reference group, blacks have higher proportions of females and immigrants, and they are less likely to be married and more likely to be unemployed and live in metropolitan areas. On the other hand, Hispanics have higher proportions of males and immigrants, and they are more likely to be married, unemployed, and live in metropolitan areas. Both blacks and Hispanics have larger family size with more own children than whites. In terms of educational attainment, blacks and Hispanics have less years of schooling relative to whites. There are also some geographical variation across groups. Blacks are more likely to live in the Midwest, Middle Atlantic, and South Atlantic regions, whereas Hispanics are more likely to live in the West and West South Central regions.

Panel B shows the descriptive statistics of the three SNAP components by subgroup. Among the analysis sample, blacks and Hispanics are substantially more likely to be eligible for SNAP. The proportions of households that are eligible are 23 percent for whites and 41 percent for both blacks and Hispanics. Among the eligible households, blacks have the highest take-up rate (62 percent), followed by whites (49 percent) and Hispanics (42 percent). Figure 2.2 also shows that blacks constantly have the highest take-up rate, followed by whites and Hispanics. Conditional on participation, Hispanics receive the highest amount of SNAP benefits on average (404 dollars per month). Blacks receive an average benefit level of 334 dollars per month, which is also significantly higher than the amount that whites receive (314 dollars per month).

Panel C displays the descriptive statistics of food-related outcomes by subgroup. With respect to food consumption, blacks spend significantly smaller amounts on food than whites, whereas Hispanics spend more than whites. With respect to food insecurity exposure, we look at two measures of food-related hardship as in Flores-Lagunes et al. (2018): incidence, the binary measure which captures whether households are food insecure; and severity, based on a continuous measure, the Rasch scale score, in the FSS. The results indicate that blacks and Hispanics suffer from food insecurity more than whites. The outcomes are similar to our previous study, except for the difference in the Rasch score between Hispanics and whites, which may be due to different sample used in the two analyses.

2.5.2 Decomposition Results

Table 2.4 presents the results of our decomposition and potential outcomes of changes in the SNAP policy rules. Row by row, we report estimates of the overall difference (Δ_Y) and the contributions from the eligibility (Δ_L), participation (Δ_T), and generosity (Δ_G) components. For the decomposition by race and ethnicity, we regard non-Hispanic whites as the reference group. We first consider the estimates from the main decomposition specified in section 3. Column (1) shows that blacks and Hispanics overall receive 32.73 and 37.48 dollars more than whites from the SNAP program per month, respectively. Using the estimated results of the MPCF out of SNAP benefits from Hastings and Shapiro (2018), SNAP increases food consumption by 16.37 to 19.64 dollars per month for blacks, and by 18.74 to 22.49 dollars per month for Hispanics, compared to whites. Differences in the proportion of being eligible alone can explain a substantial part of the overall differences in the mean benefits for both black-white and Hispanic-white differentials. In addition, differences in take-up rates lead to upward shift of the benefit levels and food consumption for blacks, but it has little impact for Hispanics. In contrast, the generosity component increases Hispanics' benefit level, but it is negligible for blacks.

To understand the magnitude of these differences in reducing food insecurity differentials, we use information on the reported resource gap to assess perceived food assistance shortcomings. There are two questions we use to capture the resource gap in the FSS: (1) how much additional money needed to meet weekly basic household food needs; (2) how much less money could be spent and still meet basic household food needs. Combining these two, we create a variable that measures the amount deviating from a default level of food spending to be food secure. Consistent with the food insecurity differentials, the average resource gaps are larger for blacks and Hispanics relative to whites, and both of the differences in average resource gaps are substantial. Black households reported having 89.97 dollars less per month than white households, and Hispanic households reported having 58.12 dollars less per month. Together with the decomposition results, SNAP reduces the food consumption gaps between blacks/Hispanics and whites by a modest amount, which is likely not enough to reduce the differences in the resource gaps between groups. However, it is important to note that there are two major caveats of this measure. First, there are a lot of missing values in the two variables used to crease the resource gap measure. Second, these self-reported amounts are subject to personal interpretation and potential mismeasurement.

In order to shed light on the extent to which SNAP may affect differences in transfer amounts across demographic groups, we then consider three scenarios that vary SNAP policy rules, similar to Section 3.2. The first scenario provides every participant the same amount of SNAP benefits, 250 dollars per month. The second scenario involves automatic enrollment, that is, all eligible households automatically enroll in the SNAP program. The third scenario is universal eligibility, which let all households become eligible for SNAP.

As seen in Table 2.4, the most effective way to reduce differences in food insecurity through SNAP is the automatic enrollment policy. Under this scenario, the SNAP benefit levels increase the most for both blacks and Hispanics relative to whites. Without the effect of differences in take-up rates, the eligibility component largely increases the benefit levels for more disadvantaged groups. On the other hand, the constant transfer policy slightly increases SNAP benefits of blacks relative to whites, but it lowers the differences in SNAP benefits between Hispanics and whites, compared to the baseline decomposition. Universal eligibility has little impact on the differences between blacks and whites, as well as Hispanics and whites. If anything, it slightly decreases the black-white differential but increases the Hispanic-white differential. Since the relative importance of the three components on differences in benefit levels are indicative of their relative importance on differences in food consumption, we would expect that the automatic enrollment policy is the most effective in alleviating differences in food consumption and exposure to food insecurity.

2.6 Conclusion

Food insecurity varies greatly by race and ethnicity, and the disparities in food insecurity have been a persistent problem in the U.S. This study attempts to uncover the pathways through which SNAP may have on the existing heterogeneity in exposure to food insecurity. We develop a sequential framework that decomposes differences in SNAP benefit levels across racial/ethnic groups into three components: differences in eligibility, participation, and generosity. We then link the results to differences in food consumption through the MPCF out of SNAP benefits to provide implications on food insecurity differentials.

Our results suggest that differences in eligibility attribute to a substantial part of the overall difference in SNAP benefit levels for both black-white and Hispanic-white differentials. On the other hand, differences in participation increase the benefit levels for blacks but have little impact for Hispanics. The generosity component is in reverse. It increases Hispanics' benefit levels but is negligible for blacks. Combining our results with the estimated MPCF out of SNAP benefits in Hastings and Shapiro (2018), SNAP reduces the differences in food consumption between blacks/Hispanics and whites by a modest amount, which is likely not enough to reduce the differences in the resource gaps between groups. To investigate potential effects of changes in SNAP policy rules on differences in benefit levels, we then carry out policy-relevant counterfactual analysis. Among the three policies under consideration, we find that automatic enrollment may be the most effective in alleviating differences in exposure to food insecurity across racial and ethnic groups. Overall, our results suggest that maintaining SNAP eligibility and increasing program take-up are critical for disadvantaged populations.

There are several future directions to take this work. First, the exploratory analysis carves out the contours of the problem, and subsequent studies can investigate the effects of more realistic policies. For instance, our model can be used to study the effects of the controversial policy proposal that raises work requirements for SNAP beneficiaries. Since this policy is likely to overwhelmingly hurt the poor who seek hunger relief through SNAP, it would be useful to know to what extent and through which pathways it may exacerbate inequality in food insecurity. In addition, there are a variety of state-based program policies that are worth exploring.[30] Our model can help investigate which policy levers can reduce the disparities more effectively. Second, future research can look at other dimensions of inequality in food insecurity exposure. For instance, Coleman-Jensen et al. (2018) finds that rates of food insecurity are substantially higher for certain populations, such as single-headed households with children and households located in South Census Region. Also, Cunnyngham (2018) shows that take-up of SNAP is disproportionately low among the elderly; in 2016, only 45 percent of eligible elderly enrolled in SNAP, compared to 85 percent overall. In the presence of potential behavioral biases, understanding how government programs mitigate the consequences of food insecurity and evaluating the welfare impact of various interventions can help policy makers orient policies more effectively and, ultimately, alleviate social and economic inequality.

[30] One example can be examining the dynamic nature of State operational policies on Broad-Based Categorical Eligibility (BBCE).

References

Alaimo, Katherine, Christine M. Olson, and Edward A. Frongillo (2001), "Food Insufficiency and American School-Aged Children's Cognitive, Academic, and Psychosocial Development," *Pediatrics*, 108 (1), 44-53.

Almond, Douglas, Hilary W. Hoynes, and Diane W. Schanzenbach (2011), "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes," *Review of Economics and Statistics*, 93 (2), 387-403.

Bellani, Luna and Michela Bia (2016), "Intergenerational Poverty Transmission in Europe: The Role of Education," Working Paper Series 2016-02, Department of Economics, University of Konstanz: University of Konstanz.

Bruich, Gregory A. (2014), "The Effect of SNAP Benefits on Expenditures: New Evidence from Scanner Data and the November 2013 Benefit Cuts," Harvard University. Mimeograph.

Coleman-Jensen, Alisha, Matthew P. Rabbitt, Christian A. Gregory, and Anita Singh (2016), "Household Food Security in the United States in 2015," Economic Research Report 215, U.S. Department of Agriculture Economic Research Service.

Coleman-Jensen, Alisha, Matthew P. Rabbitt, Christian A. Gregory, and Anita Singh (2018), "Household Food Security in the United States in 2017," Economic Research Report 256, U.S. Department of Agriculture Economic Research Service.

Cunnyngham, Karen (2018), "Trends in Supplemental Nutrition Assistance Program Participation Rates: Fiscal Year 2010 to Fiscal Year 2016," U.S. Department of Agriculture Food and Nutrition Service.

Currie, Janet M. (2003), "U.S. Food and Nutrition Programs," in *Means-Tested Transfer Programs in the United States*, NBER Conference Report Series, ed. Robert A. Moffitt, Chicago and London: University of Chicago Press, 199-289.

--- (2006), *The Invisible Safety Net: Protecting the Nation's Poor Children and Families*, Princeton and Oxford: Princeton University Press.

--- (2009), "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature*, 47 (1), 87-122.

Flores-Lagunes, Alfonso, Hugo Jales, Judith Liu, and Norbert L. Wilson (2018), "The Differential Incidence and Severity of Food Insecurity by Racial, Ethnic, and Immigrant Groups over the Great Recession in the United States," *American Economic Review: Papers & Proceedings*, 108 (3), 379-83.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo (2011), "Decomposition Methods in Economics," in *Handbook of Labor Economics*, Vol. 4, Part A, ed. Ashenfelter Orley and Card David: Elsevier, 1-102.

Gundersen C. and B. Kreider (2008), "Food Stamps and Food Insecurity: What Can Be Learned in the Presence of Nonclassical Measurement Error?" *Journal of Human Resources*, 43(2): 352-382.

Gundersen, Craig, Brent Kreider, and John Pepper (2011), “The Economics of Food Insecurity in the United States,” *Applied Economic Perspectives and Policy*, 33 (3), 281-303.

Gundersen, Craig, Brent Kreider and John V. Pepper (2018), “Reconstructing the Supplemental Nutrition Assistance Program to More Effectively Alleviate Food Insecurity in the United States,” *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 4 (2), 113-130.

Hastings, J. and J. M. Shapiro (2018), “How Are SNAP Benefits Spent? Evidence from a Retail Panel.” *American Economic Review*, 108(12): 3493–3540.

Hoynes, H. W. and D. W. Schanzenbach (2009), “Consumption Responses to In-Kind Transfers: Evidence from the introduction of the Food Stamp Program,” *American Economic Journal: Applied Economics*, 1(4): 109-139.

Jensen, Helen H. (2002), “Food Insecurity and the Food Stamp Program,” *American Journal of Agricultural Economics*, 84 (5), 1215-28.

Montez, Jennifer Karas, Anna Zajacova, Mark D. Hayward, Steven H. Woolf, Derek Chapman, Jason Beckfield, “Educational Disparities in Adult Mortality Across U.S. States: How Do They Differ, and Have They Changed Since the Mid-1980s?” *Demography*, forthcoming.

Nam, Yunju, Jin Huang, Colleen Heflin, and Michael Sherraden (2015), “Racial and Ethnic Disparities in Food Insufficiency: Evidence from a Statewide Probability Sample,” *Journal of the Society for Social Work and Research*, 6 (2), 201-28.

Ratcliffe, Caroline, Signe-Mary McKernan, and Sisi Zhang (2011), “How Much Does the Supplemental Nutrition Assistance Program Reduce Food Insecurity?” *American Journal of Agricultural Economics*, 93 (4), 1082-98.

Ratcliffe, Caroline (2015), “Child Poverty and Adult Success,” in *Brief*, Washington, DC: The Urban Institute.

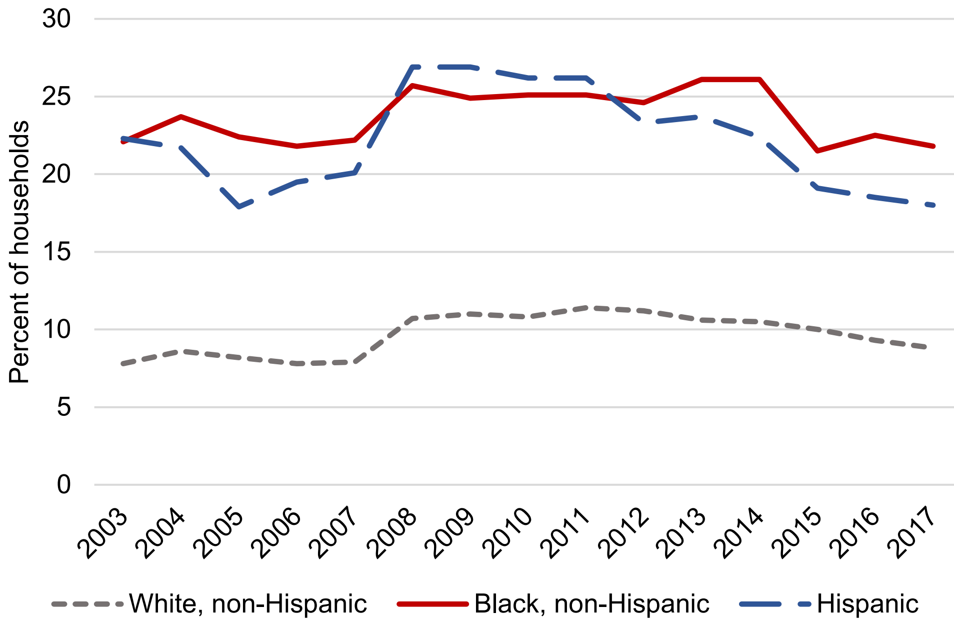
Schanzenbach, Diane W., Lauren Bauer, and Greg Nantz (2016), “Twelve Facts About Food Insecurity and SNAP,” in *The Hamilton Project*, Washington, DC: Brookings Institution.

U.S. Department of Agriculture (2019), “Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2017,” U.S. Department of Agriculture, Food and Nutrition Service.

Wilde, P. and M. Nord (2005), “The Effect of Food Stamps on Food Security: A Panel Data Approach,” *Review of Agricultural Economics* 27 (3): 425-432.

Wolkwitz, K. (2008), “Trends in Food Stamp Program Participation Rates: 2000 to 2006,” U.S. Department of Agriculture, Food and Nutrition Service.

Figure 2.1: Food Insecurity by Race/Ethnicity, 2003-2017



Source: USDA, Economic Research Service.

Figure 2.2: SNAP Participation Rates, 2003-2014

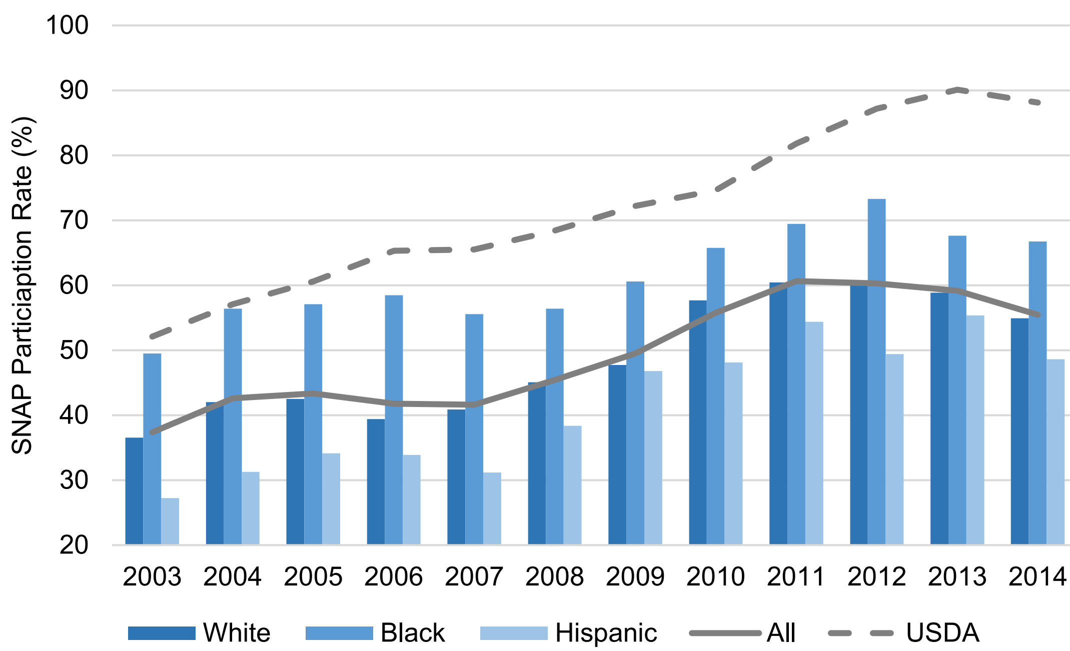


Table 2.1: Setup of the Numerical Example of the SNAP Decomposition

	$Pr(L_g = 1)$	$Pr(T_g = 1 \mid L_g = 1)$	$Pr(H_g = 1 \mid L_g = 1, T_g = 1)$
$g = 1$ (blacks)	0.5	0.5	0.3
$g = 0$ (whites)	0.3	0.6	0.25

Table 2.2: Results of the Numerical Example of SNAP Decomposition

	Baseline (1)	Constant Transfer (2)	Auto Enrollment (3)	Universal Eligibility (4)
Overall difference	25	35	68.75	-25
Decomposition:				
Eligibility	32.5	50	65	0
Participation	-9.75	-15	0	-32.5
Generosity	2.25	0	3.75	7.5

Table 2.3: Descriptive Statistics by Race and Ethnicity

A. Socioeconomic Characteristics

	White, non-Hispanic	Black, non-Hispanic	Hispanic	Difference: Black-White	Difference: Hispanic-White
Monthly Earnings	2311.088 (5.984)	2031.322 (11.459)	1861.316 (8.195)	-279.766*** (14.616)	-449.772*** (12.076)
Age	38.006 (0.046)	37.902 (0.101)	35.121 (0.075)	-0.104 (0.115)	-2.885*** (0.096)
Male	0.480 (0.002)	0.402 (0.004)	0.549 (0.003)	-0.078*** (0.004)	0.069*** (0.004)
Immigrant	0.037 (0.001)	0.116 (0.002)	0.569 (0.003)	0.079*** (0.003)	0.532*** (0.003)
Married	0.473 (0.002)	0.317 (0.004)	0.500 (0.003)	-0.156*** (0.004)	0.027*** (0.004)
Unemployed	0.031 (0.001)	0.047 (0.002)	0.040 (0.001)	0.016*** (0.002)	0.009*** (0.001)
Household Head	0.519 (0.002)	0.581 (0.004)	0.456 (0.003)	0.062*** (0.004)	-0.063*** (0.004)
Number of Own Children	0.923 (0.004)	1.062 (0.010)	1.277 (0.008)	0.139*** (0.010)	0.354*** (0.009)
Family Size	2.962 (0.005)	3.132 (0.013)	3.845 (0.012)	0.170*** (0.013)	0.883*** (0.012)
Metropolitan Area	0.698 (0.002)	0.858 (0.003)	0.888 (0.002)	0.159*** (0.003)	0.189*** (0.002)
Education					
12 grades or less	0.112 (0.001)	0.152 (0.003)	0.401 (0.003)	0.040*** (0.003)	0.288*** (0.003)
High school degree	0.348 (0.002)	0.378 (0.004)	0.316 (0.003)	0.029*** (0.004)	-0.032*** (0.003)
Some college or Associate's degree	0.334 (0.002)	0.329 (0.004)	0.207 (0.003)	-0.005 (0.004)	-0.127*** (0.003)
Bachelor's degree	0.152 (0.001)	0.101 (0.002)	0.061 (0.001)	-0.050*** (0.003)	-0.091*** (0.002)
Master's degree or above	0.053 (0.001)	0.039 (0.001)	0.015 (0.001)	-0.014*** (0.002)	-0.038*** (0.001)
Family Income					
Less than $10,000	0.075 (0.001)	0.135 (0.003)	0.092 (0.002)	0.060*** (0.003)	0.018*** (0.002)
$10,000 to $19,999	0.146 (0.001)	0.202 (0.003)	0.205 (0.003)	0.055*** (0.003)	0.059*** (0.003)
$20,000 to $29,999	0.187 (0.001)	0.214 (0.003)	0.236 (0.003)	0.027*** (0.003)	0.049*** (0.003)
$30,000 to $39,999	0.152 (0.001)	0.150 (0.003)	0.189 (0.002)	-0.002 (0.003)	0.037*** (0.003)

\$40,000 to \$49,999	0.106 (0.001)	0.084 (0.002)	0.100 (0.002)	-0.022*** (0.002)	-0.006*** (0.002)
\$50,000 to \$59,999	0.085 (0.001)	0.061 (0.002)	0.060 (0.001)	-0.024*** (0.002)	-0.025*** (0.002)
\$60,000 to \$74,999	0.090 (0.001)	0.060 (0.002)	0.048 (0.001)	-0.030*** (0.0020	-0.041*** (0.002)
\$75,000 to \$99,999	0.078 (0.001)	0.039 (0.001)	0.035 (0.001)	-0.039*** (0.002)	-0.043*** (0.001)
\$100,000 to \$149,999	0.048 (0.001)	0.026 (0.001)	0.018 (0.001)	-0.022*** (0.001)	-0.030*** (0.001)
\$150,000 or more	0.017 (0.000)	0.009 (0.001)	0.005 (0.000)	-0.007*** (0.001)	-0.011*** (0.001)
Census Region					
New England	0.120 (0.001)	0.037 (0.001)	0.042 (0.001)	-0.084*** (0.002)	-0.079*** (0.002)
Middle Atlantic	0.076 (0.001)	0.087 (0.002)	0.075 (0.002)	0.011*** (0.002)	-0.002 (0.002)
East North Central	0.136 (0.001)	0.114 (0.002)	0.060 (0.001)	-0.021*** (0.003)	-0.076*** (0.002)
West North Central	0.166 (0.001)	0.063 (0.002)	0.050 (0.001)	-0.103*** (0.002)	-0.116*** (0.002)
South Atlantic	0.141 (0.001)	0.374 (0.004)	0.137 (0.002)	0.234*** (0.004)	-0.004 (0.002)
East South Central	0.060 (0.001)	0.103 (0.002)	0.013 (0.001)	0.043*** (0.002)	-0.047*** (0.001)
West South Central	0.077 (0.001)	0.139 (0.003)	0.184 (0.002)	0.062*** (0.003)	0.107*** (0.003)
Mountain	0.129 (0.001)	0.031 (0.001)	0.165 (0.002)	-0.098*** (0.002)	0.036*** (0.003)
Pacific	0.094 (0.001)	0.051 (0.002)	0.275 (0.003)	-0.043*** (0.002)	0.181*** (0.003)
Observations	90,924	17,255	25,662		

B. SNAP Components

	White	Black	Hispanic	Difference: Black-White	Difference: Hispanic-White
Eligibility	0.234 (0.001)	0.410 (0.004)	0.411 (0.003)	0.176*** (0.004)	0.177*** (0.003)
Conditional on eligibility					
Take-up	0.487 (0.003)	0.617 (0.006)	0.424 (0.005)	0.130*** (0.007)	-0.063*** (0.006)
Conditional on take-up					
Benefits (self-reported)	257.738 (1.580)	260.137 (2.393)	272.719 (2.324)	2.399 (2.894)	14.982*** (2.850)
Benefits (imputed)	314.172 (2.562)	334.005 (4.080)	404.200 (4.178)	19.832*** (4.718)	90.028*** (4.680)
Cash-on-hand for food relative to basic food needs					
Total food cash	1.260 (0.149)	-21.232 (0.434)	-13.269 (0.328)	-22.492*** (0.394)	-14.529*** (0.331)
Food cash without SNAP benefits	-25.981 (0.371)	-83.526 (1.196)	-58.403 (0.863)	-57.545*** (1.006)	-32.422*** (0.837)

C. Food-related Outcomes

	White	Black	Hispanic	Difference: Black-White	Difference: Hispanic-White
Food consumption					
Total expenditures on food last week	135.310 (0.331)	118.253 (0.779)	140.004 (0.649)	-17.058*** (0.837)	4.694*** (0.714)
Usual expenditures on food per week	120.280 (0.265)	108.271 (0.638)	130.322 (0.554)	-12.009*** (0.672)	10.041*** (0.580)
Food insecurity					
Binary indicator	0.259 (0.001)	0.383 (0.004)	0.353 (0.003)	0.124*** (0.004)	0.094*** (0.003)
Rasch score	4.221 (0.012)	4.519 (0.024)	4.271 (0.019)	0.297*** (0.027)	0.050** (0.023)

Note: This table reports means and standard deviations (in parentheses) for the analysis samples in the CPS-FSS between 2003 and 2016. We focus on households with incomes below 185 percent of the poverty line or that report being short of money for food (the target population of the FSS).

Table 2.4: Estimated Results of the SNAP Decomposition

	Baseline (1)	Constant Transfer (2)	Auto Enrollment (3)	Universal Eligibility (4)
A. Black-White				
Overall difference	32.733	35.362	51.392	31.730
Decomposition:				
Eligibility	26.315	21.992	68.584	0
Participation	8.072	13.370	0	-1.861
Generosity	-1.654	0	-17.192	33.592
B. Hispanic-White				
Overall difference	37.477	16.238	88.443	39.447
Decomposition:				
Eligibility	27.760	11.694	88.615	0
Participation	2.229	4.544	0	-41.379
Generosity	7.487	0	-0.171	80.826

Chapter 3

The Effects of Trade Liberalization and Domestic Reforms on Income Inequality in a Middle-Income Open Economy: Evidence from Taiwan

Mei-Ying Lai
Department of Accounting Information
Tainan University of Technology, Tainan, Taiwan
t20107@mail.tut.edu.tw

Judith Liu
Department of Economics and Center for Policy Research
Syracuse University, Syracuse, New York, 13244-1020
jliu61@syr.edu

Zong-Shin Liu
Department of Economics
Feng Chia University, Taichung, Taiwan
zsliu@fcu.edu.tw

3.1 Introduction

Since the 1980s, income inequality has increased in most countries, and over the same time period, trade liberalization has taken place rapidly. A natural question is whether trade liberalization leads to the rising inequality. According to the Stolper-Samuelson (S-S) theorem (Stolper and Samuelson, 1941), international trade raises the relative wage of skilled workers and deteriorates income inequality in developed countries, whereas it increases the relative wage of unskilled workers and improves income inequality in developing countries. However, the estimation results are mixed either using a cross-country model (e.g., Savvides, 1998; Barro, 2000; Reuvny and Li, 2003; Milanovic, 2005; IMF, 2007; Dreher and Gaoton, 2008) or an individual country model (e.g., Beyer et al., 1999; Chen and Hsu, 2001; Galiani and Sanguinetti, 2003; Mah, 2003; Herault, 2007; Kumar and Mishra, 2008; Sato and Fukushige, 2009; Mcnabb and Said, 2013; Lai et al., 2019), and thus the S-S theorem does not seem to be supported by empirical evidence.

Rather than trade liberalization, some studies show that domestic reforms (such as technological progress and financial liberalization) might be the main driving force of income inequality, particularly in developing countries (e.g., Goldberg and Pavcnik, 2007; IMF, 2007). Previous research has explored the effects of economic freedom on income inequality using a cross-country model, where economic freedom can be measured by a composite index constructed with many policy components, such as international trade, government regulations, and taxation (Berggren, 1999; Carter, 2006). However, these empirical results are also unclear (e.g., Berggren, 1999; Scully, 2002; Carter, 2006; Bergh and Nilsson, 2010). It is worth noting that the findings using a cross-country model may not be applicable to an individual country. Furthermore, using the composite index to measure economic freedom cannot reveal the impact of a specific policy on income inequality (Carter, 2006). Thus, a more thorough policy analysis in an individual country setting is needed.

This paper aims to estimate the effects of trade liberalization on household income inequality and to investigate whether trade liberalization or domestic reforms are the leading factors of the rising inequality since 1980 in Taiwan. Following the globalization trends, Taiwan adopted trade liberalization policy in the early 1980s. Since then, the degree of trade openness (the share of exports plus imports to GDP) has substantially increased from an average of 85.4% in the 1980s to above 100% after the 2000s. Trade expansion has led to economic growth in Taiwan. For example, the average income per capita was only US$4,129.9 in the 1980s, but it exceeded US$20,000 in 2011 and reached US$23,131 in 2015.[31] On the other hand, household income inequality has also increased over time. For example, the quintile ratio of household disposable income, which is the relative disposable income of the highest 20% to the lowest 20% income households, rose from 4.2 in 1980 to above 6.0 after 2001.

In addition to external trade liberalization, the Taiwanese government has implemented various domestic reforms since 1980. We focus on four of the reforms that may influence income inequality: (i) financial liberalization in the 1980s and two financial reforms in the 2000s; (ii) the expansion of higher education since 1985; (iii) the amendments to industrial development policies that encourage investment in capital assets and research and development (R&D); (iv) large increases in social welfare and social insurance expenditures since the 1990s. The goals of these reforms were to attain a better functioning economy, to stimulate economic growth, or to meet political needs. Nevertheless, they may also have affected income inequality and had a

[31] National Development Council, R.O.C. (Taiwan), *Taiwan Statistical Data Book*, 2016.

larger impact than trade liberalization (Behrman et al., 2003; IMF, 2007; Goldberg and Pavcnik, 2007).

Based on the Heckscher-Ohlin (H-O) model, previous studies using a cross-country model all dichotomize the countries of interest into developed and developing countries. However, many economies, such as Taiwan and some OECD countries (e.g., Mexico, Turkey, Chile, Latvia, Greece, Poland, and Hungary), should be considered as a middle-income open economy (MIOE). They simultaneously trade with both more- and less-developed countries, and the trade effects from one side may offset the other, resulting in a small or even insignificant overall effect on income inequality. Compared to many advanced OECD countries, Taiwan is a MIOE in terms of its income per capita.[32] Departing from the conventional approach, Chen and Hsu (2001) regard Taiwan as a MIOE and separately estimate the effects of net exports to OECD countries and those to non-OECD countries on wage differentials in Taiwan. To provide a complete picture of the trade effects on inequality in a MIOE, we also distinguish the effects of net exports between OECD and non-OECD countries.

We use the quintile ratio of household disposable income to measure income inequality. To assess the effects of trade liberalization and domestic reforms on inequality, we construct an empirical model based on decomposition of the sources of household disposable income in the quintile ratio to capture possible influence factors. The model includes factors of trade liberalization and domestic reforms as explanatory variables, and thus it allows us to analyze the relative importance of each factor.

[32] The income per capita in terms of 2015 US dollars is 23,131 in Taiwan, which is substantially smaller than some OECD countries, such as the US (56,070), Germany (45,780), United Kingdom (43,700), France (40,530), and Japan (38,780), according to *Taiwan Statistical Data Book* (2016) and *World Development Indicators*, World Bank.

Using time-series data from 1980 to 2015, we conduct a cointegration test to estimate the effects of various influence factors on income inequality. Our main findings are as follows. First, net exports to OECD countries significantly increase inequality, whereas net exports to non-OECD countries insignificantly decrease inequality. Overall, trade liberalization increases inequality in Taiwan. Second, among domestic reforms, government expenditures on social welfare and social insurance reduce inequality. On the other hand, technological progress biased toward skilled labor, industrial structural change, and financial market reforms increase inequality. Third, by calculating the long-run mean impacts to evaluate the contribution of each influence factor, we find that technological progress and industrial structural change are the main driving forces of the rising inequality in Taiwan.

This paper contributes to the literature in several ways. First, we jointly estimate the impacts of trade liberalization and domestic reforms in an attempt to explain the rising trend in income inequality in Taiwan. This approach better identifies the influence factors of income inequality in an individual country setting. Second, we investigate the case of Taiwan to provide some empirical evidence on the effects of trade with more- and less-developed countries in a MIOE. Third, we construct an empirical model to assess possible influence factors of household income inequality, and this framework can be applied to other studies on inequality. Fourth, our results provide policy implications for some countries, particularly developing countries, which endeavor in various reforms to promote economic growth but at the same time suffer from the deterioration of income distribution.

This paper is organized as follows. Section 2 discusses the findings in the literature. Section 3 provides background on income inequality, policy changes in trade liberalization, and domestic reforms in Taiwan. Section 4 demonstrates the empirical model based on quintile ratio decomposition, and Section 5 presents the results. Section 6 concludes and discusses the implications of these results.

3.2 Literature Review

The two main theories used in the literature to study the effects of international trade on income distribution are the H-O theory and the S-S theorem. According to the H-O theory and following the comparative advantage principle, developed countries, which are relatively abundant in skilled labor, export skilled-labor-intensive products to developing countries and import unskilled-labor-intensive products from them. On the other hand, developing countries, which are relatively abundant in unskilled labor, export unskilled-labor-intensive products to developed countries and import skilled-labor-intensive products from them. By the S-S theorem, an expansion in net exports in a developed (developing) country increases the relative demand for skilled (unskilled) labor, which increases (decreases) the relative wage of skilled labor. This in turn causes income inequality to rise (fall) in a developed (developing) country, *ceteris paribus*.

A number of studies have estimated the effects of trade liberalization on income inequality, but the results are mixed, with some being inconsistent with the prediction of the S-S theorem. For instance, using a cross-country model, Barro (2000) and Milanovic (2005) find that trade openness decreases inequality in high-income countries but increases inequality in low-income countries. Savvides (1998) shows that trade openness raises inequality in less-developed countries, while the effect is statistically insignificant in developed countries. Reuveny and Li (2003) and IMF (2007) demonstrate that economic globalization improves inequality in both developed and developing countries. In addition to economic globalization, Dreher and Gaoton

(2008) also consider social and political globalization. They find that globalization exacerbates inequality, particularly in OECD countries, but there is no consistent effect in less-developed countries.

The empirical results using an individual country model are also mixed. Wood (1997) finds that trade opening in the 1960s and 1970s in East Asian countries decreases wage differentials between skilled and unskilled workers, but trade opening in the 1980s in Latin American countries increases wage differentials. Several studies also show that trade openness widens wage differentials in Latin American countries, e.g., Chile (Beyer et al., 1999), Mexico (Hanson and Harrison, 1999), and Argentina (Galiani and Sanguinetti, 2003). In terms of Asian countries, Mah (2003) finds that trade liberalization does not have a significant impact on income inequality in Korea, but Sato and Fukushige (2009) show that economic globalization reduces income inequality in Korea in both the short run and the long run. Kumar and Mishra (2008) investigate the 1991 trade liberalization reforms in India, and Mcnabb and Said (2013) examine trade liberalization policy since the mid-1980s in Malaysia. Both studies find that trade liberalization decreases wage differentials between skilled and unskilled workers. As pointed out by Goldberg and Pavcnik (2007), there is limited evidence that supports the conventional S-S theorem based on previous studies on the trade effects in developing countries.

In the case of Taiwan, Chan et al. (1999) find a positive and significant effect of net exports on wage differentials between skilled and unskilled labor. When distinguishing the effects among trade partners, Chen and Hsu (2001) show that net exports to OECD countries increase wage differentials, whereas net exports to non-OECD countries decrease wage differentials. On the other hand, Lai et al. (2019) suggest that trade openness leads to a decrease in income inequality overall.

Regarding the effects of economic liberalization policies on income inequality, the empirical results are also inconclusive. Using a cross-country model and a composite index to measure economic freedom, Berggren (1999) finds a positive relationship between economic freedom and equality, which is mainly attributed to trade liberalization and financial deregulation. Scully (2002) also shows that economic freedom promotes income inequality. In contrast, Carter (2006) and Bergh and Nilsson (2010) suggest that economic freedom increases inequality, especially in rich countries. Behrman et al. (2003) investigate the effects of six economic policy changes on wage differentials in 18 Latin American countries and find that liberalization policy changes together have a disequalizing effect in the short run. This is due to the impact of financial market reform, capital account liberalization, and tax reform; however, trade openness has no significant effect on wage differentials. As mentioned earlier, the results of a cross-country model may not be applicable to an individual country, and the usage of the composite index to measure economic freedom cannot reveal the effect of a specific reform on inequality.

There are several reasons why the effects of trade on income inequality are inconsistent or statistically insignificant in the literature. First, previous studies do not distinguish the trade effects between different types of trade partners. In fact, many countries are MIOEs trading with both developed and developing countries, and the effects from one side may conflict and offset the other, resulting in a small or even insignificant total effect on income inequality. Since Taiwan is a MIOE, we split Taiwan's trade partners into OECD and non-OECD countries and separately estimate their effects on income inequality, similar to Chen and Hsu (2001). Second, the mixed results could be due to differences in the selection of domestic (control) variables.

Nevertheless, there is no standard empirical model from which domestic variables could be drawn (Carter, 2006).

For the purpose of this study, we decompose the quintile ratio based on the sources of household disposable income to directly capture possible influence factors of inequality. Instead of using ordinary least squares (OLS) estimation, as most of the literature does, we conduct a cointegration test to estimate the long-run effects of the influence factors on income inequality. Since trade liberalization and domestic reforms gradually take place, the effects on inequality take time to develop and may persist for a long period of time. It is more applicable to estimate the long-run impact with time-series data. Moreover, from the econometric point of view, the variables considered in our model are all integrated of degree one, I(1). Using OLS estimation with the first-differenced values may suffer from over-differencing problem and result in biased estimates. With the cointegration test, we can obtain reliable estimates of the long-run effects.

3.3 Income Inequality, Trade Liberalization, and Domestic Reforms

3.3.1 Household Income Inequality

We use data from the *Report on the Survey of Family Income and Expenditure* (*RSFIE*), issued annually by the Directorate-General of Budget, Accounting and Statistics (DGBAS), Executive Yuan of Taiwan. The *RSFIE* is the only official data source of family income and expenditure, as well as income inequality.[33] There are two main indicators of income inequality reported by the DGBAS: the quintile ratio and the Gini coefficient. As shown in Figure 3.1, these two indicators have very similar trends. The quintile ratio was 4.17 in 1980, reached the peak of

[33] The DGBAS conducts a survey each year to obtain family income and expenditure from 13,600-16,400 families in Taiwan. This survey started from 1964, but it was carried out on a two-year basis prior to 1972. The sampling rate is about 2-4‰ of the entire population.

6.39 in 2001, and remained stable at around 6.10 afterwards. Similarly, the Gini coefficient was 0.28 in 1980, reached the peak of 0.35 in 2001, and then remained stable at around 0.34.

We use the quintile ratio of household disposable income to measure income inequality since it is the most commonly used indicator in Taiwan. The quintile ratio is defined as the ratio of disposable income of the highest 20% to the lowest 20% income households. We consider inequality in disposable income at the household level because government transfer payments on social welfare and social insurance in Taiwan are based on household income. As will be discussed later, those payments are relevant to inequality through income redistribution effects. By decomposing the sources of household disposable income in the quintile ratio, we construct an empirical model that captures possible influence factors of income inequality.

Besides trade liberalization and domestic reforms, differences in household characteristics may also affect income inequality. Table 3.1 shows three differences in household characteristics between the highest 20% income (the fifth quintile) and the lowest 20% income (the first quintile) households: the number of persons per household, the number of persons employed per household, and the educational attainment of economic household heads. Along with economic development and social structural change, the average number of persons per household and the average number of persons employed per household have gradually decreased in both the first and fifth quintiles since 1980. However, the ratio of the number of persons employed per household of the fifth quintile to that of the first quintile has increased from 1.82 in 1980 to 5.04 in 2015. As the difference in the number of persons employed between the highest- and lowest-income households becomes larger, the household income differential may also increase, causing income inequality to rise.[34]

[34] In Section 4, we provide more details about the influence of the number of persons employed per household on income inequality.

Regarding the percentage of economic household heads with a bachelor's degree or higher, the fifth quintile always has a higher rate than the first quintile. Typically, higher educational attainment leads to higher earnings, and thus household income of the fifth quintile should be higher. Nevertheless, the massive expansion in higher education in Taiwan since the 1990s has decreased the difference in educational attainment between the highest- and lowest-income households. As shown in Table 3.1, the percentages of economic household head with a bachelor's degree or higher in both the first and the fifth quintile households have risen rapidly, but the ratio of the two percentages has gradually declined from 16.20 in 1980 to 7.89 in 2015. This implies that the expansion of higher education may reduce household income inequality. In Table 3.2, we summarize changes in trade liberalization policy and four domestic reforms in Taiwan since the 1980s, and further discuss them in the following subsections.

3.3.2 Trade Liberalization

During the early stages of economic development, trade policy in Taiwan was associated with foreign exchange control policy due to the shortage in foreign exchange. In the early 1980s, Taiwan's trade surplus substantially increased, and a large amount of foreign exchange reserves were accumulated. As a result, the major trade partners, particularly the U.S., required Taiwan to reduce trade barriers and appreciate the New Taiwan (NT) dollar. The Taiwanese government subsequently implemented trade liberalization policy.

This policy includes two main components. First, it relaxed foreign exchange control in 1987 and adjusted the exchange rate system from a fixed to a managed floating system in 1989. Second, the policy reduced the import tariff and relaxed controls on exports and imports. Since the mid-1980s, Taiwan started to lower its import tariff rate (tariff revenues/total value of imports) from an average of 7.5%. In 2002, Taiwan became a member of the World Trade Organization (WTO), and the average tariff rate substantially declined to the current level of around 1.5%. In addition, the bilateral trade with China has greatly expanded since the 1990s. Taiwan allowed indirect trade with China in 1993,[35] and it further allowed direct trade in 2002.

As shown in Figure 3.2, the ratio of total net exports (NX) to GDP (Y) has a hump-shape. It increased from about zero to a peak of 20% in 1986 and then declined to below 5% in the 1990s. It slightly increased again after 2010. This trend indicates that trade liberalization in the mid-1980s led to a more rapid growth in imports than in exports. Since Taiwan is a MIOE, we split the total net exports into net exports to OECD countries and those to non-OECD countries. The ratio of net exports to OECD countries to GDP (NX_O / Y) dramatically decreased after the mid-1980s and became negative (trade deficit) after 1992. In contrast, the ratio of net exports to non-OECD countries to GDP (NX_{NO} / Y) remained positive (trade surplus) and increased over time. The increase in NX_{NO} / Y was relatively large in the early 1990s and 2000s, which is mainly due to trade with China. As will be discussed later, NX_O / Y and NX_{NO} / Y have different effects on income inequality.

3.3.3 Industrial Development Policy: Structural Change and Technological Progress

In the process of economic development, the Taiwanese government endeavored to promote industrial development via legislation. There were three industrial statutes: (i) Statute for the Encouragement of Investment (SEI) from 1960 to 1990, (ii) Statute for Upgrading Industry (SUI) from 1991 to 2009, and (iii) Statute for Industrial Innovation (SII) after 2010.

[35] Due to the special relationship across the Taiwan straits, Taiwanese firms were only allowed to conduct transit trade with China through Hong Kong.

These three statutes, the so-called "Industrial Constitution" in Taiwan, dominated industrial investment and contributed to technological progress.

Tax encouragement was the main policy tool of these three statutes, in particular the five-year exemption of profit-seeking enterprises income tax for the industries or firms that met the encouragement requirements. However, these three statutes targeted different types of industries and firms. The earliest statute, SEI, focused on promoting capital investment of productive industries. It implemented encouragement on specific industries and adjusted the categories of the encouraged industries based on the stage of economic development. For example, it increased encouragement for investment in the newly strategic industries (information, electricity, and machinery) in the 1980s. The second statute, SUI, aimed at upgrading industries; therefore, it changed from the previous "industry species" encouragement to "function species" encouragement, and only a few former enterprises were still included.[36] The function species encouragement included investment expenditures in automatic equipment, R&D, and labor force development. The encouragement applied to all enterprises, including small and medium enterprises, and in particular the service business. The third statute, SII, was designed to enhance industrial innovation and continue the function species encouragement. However, only R&D expenditure remained on the list, and all the industry species encouragement was eliminated. The SII applied to all enterprises, businesses, and inviable assets.

The tax encouragement of these three statutes led to two apparent changes in industrial development in Taiwan. The first one includes changes in industrial structure and employment structure. In the 1960s and 1970s, the industrial sector grew rapidly, and beginning in the 1980s

[36] For example, several important technology enterprises (newly formed industries such as communication, information, semiconductor, etc.) and the venture capital enterprise were involved in 1990, but only the newly strategic industries remained in 2000.

it moved toward high-technology manufacture. After the late 1980s, the service sector began to thrive and surpass the industrial sector. As of 2015, the share of the product value in GDP was 63.17% in the service sector and 35.13% in the industrial sector. Because the service sector absorbed more employees, changes in industrial structure also led to structural change in the labor market. As shown in Figure 3.3a, the number of employees in the industrial sector (EMP_{IND}) was larger than that in the service sector (EMP_{SER}) before 1988. Since then, EMP_{SER} has exceeded EMP_{IND}, and the ratio of EMP_{SER}/EMP_{IND} has increased substantially. The percentage of the number of sectoral employees to the number of total employees in 2015 was 59.02% in the service sector and 36.03% in the industrial sector, resulting in a ratio (EMP_{SER}/EMP_{IND}) of 1.64. Noticeably, EMP_{SER}/EMP_{IND} has remained stable at around 1.64 since 2002, implying that the industrial structure and the distribution of the number of employees between these two sectors have become stable.

The second apparent change includes rapid capital accumulation and technological progress. The earliest SEI was helpful for accumulating capital, and the function species encouragement of SUI and SII since 1991 stimulated R&D investment. All three statutes enhanced industrial technology progress. Following Lawrence and Slaughter (1993), Chan et al. (1999), Chen and Hsu (2001), Mcnabb and Said (2013), and Lai et al. (2019), we use the total factor productivity index (TFP) in the service and industrial sectors to measure technological progress.[37] As shown in Figure 3.3b, TFP has constantly increased since the 1980s.

[37] Several measures of technological change have been used in the literature. For instance, besides the total factor productivity (TFP), others also use the share of information and communication technology capital (K_{ICT}) in the total capital stock (IMF, 2007), the ratio of K_{ICT} flow to GDP (Asteriou et al., 2014), and the ratio of expenditure on R&D to sales (Berman et al., 1994).

3.3.4 Reforms in Higher Education

Over the past five decades, the higher education system in Taiwan has undergone substantial changes. In 1968, the compulsory education was extended from 6 years (primary school) to 9 years (junior high school). According to the *Education Statistics* from the Ministry of Education (MOE), the net enrollment rate of junior high school for the population aged 12-14 was 97.82% in the 2015 academic year. Traditionally, the MOE imposed strict restrictions on the establishment of new universities and college enrollment in order to maintain the quality of higher education. Driven by the need for more-educated workers to help economic development and by people's desire to attend higher education, the MOE relaxed the restrictions on establishment of universities in 1985.[38] The number of higher education institutions has begun to increase since 1987. The total number of universities and colleges increased from 28 in 1985 to 127 in 2000 and 158 in 2015. In the 1990s, the number of people who have a bachelor's degree or higher also increased substantially, from 191,752 in the 1985 academic year to 647,920 in the 2000 academic year. It further increased to 1,332,245 in the 2015 academic year, in which the net enrollment rate of universities and colleges for the population aged 18-21 was 70.86%.

We define employees aged 15 and above who have a bachelor's degree or higher (higher education) as skilled workers (LS_C) and those with a junior high school degree or lower (mandatory education) as unskilled workers (LS_M).[39] We use the ratio of LS_C / LS_M to represent the relative supply of skilled and unskilled workers. As shown in Figure 3.3c, LS_C / LS_M was only 0.15 in 1980. It increased rapidly in the 1990s, and LS_C began to exceed LS_M in 2004, with

[38] See Gindling and Sun (2002) for a detailed description of higher education planning in Taiwan.

[39] The minimum working age is 15 in Taiwan. Following Chan et al. (1999) and Lai et al. (2019), we do not consider the employees with a senior high school degree (both general and vocational) because workers with a senior high school degree may work in high-skilled, mid-skilled, and low-skilled occupations. We are unable to identify who are skilled and who are unskilled labor within this group.

LS_C / LS_M being 1.08. Since 2011, LS_C / LS_M has been larger than 2. Overall, the expansion of higher education in Taiwan increased the relative supply of skilled labor since the mid-1990s.

3.3.5 Financial Market Reforms

The financial market in Taiwan was under strict control until the government adopted liberalization and globalization policies in the 1980s. The key policy changes were (i) allowing domestic banks to open more branches and allowing foreign banks to set branches in 1984, (ii) liberalizing bank interest rates and allowing banks to determine their own interest rates for deposits and loans in 1989, and (iii) approving 15 new banks to set up in 1991.[40] However, these policies resulted in violent competition among banks and deteriorated the earnings ratio and non-performing loan ratio of banks.

To improve the financial quality of banks, the government implemented the first financial reform in 2002 and proposed the "two-five-eight plan," i.e., lowering the non-performing loan ratio of banks to below 5% within two years and maintaining the capital adequacy ratio of banks at above 8%. Since the number of existing banks was still large after the 2002 reform, the second financial reform was implemented in 2004 in order to increase the earnings ratio by reducing the number of banks. The main objectives were to (i) decrease the number of public-owned financial institutions from 12 to 6 by the end of 2005 and (ii) reduce the number of financial holdings corporations from 14 to 7 by the end of 2006. Although the number of domestic banks has diminished, these objectives have not been completely achieved.

Following IMF (2007) and Lai et al. (2019), we measure the degree of financial development using the ratio of credit to the private sector provided by commercial banks and

[40] After that, the number of banks largely increased.

other financial institutions ($CREDIT$) to GDP ($CREDIT/Y$). As shown in Figure 3.3d, financial liberalization largely expanded private credit in the 1980s, but $CREDIT/Y$ was contracted in the mid-1990s. It then increased again after the 2002 and 2004 financial reforms.

3.3.6 Changes in Social Welfare and Social Insurance policies

The transfer payments between the government and households include three parts: taxes, social welfare, and social insurance, which all have income redistribution effects. Before 1990, the majority of transfer payments in Taiwan were (i) individual income tax and (ii) compulsory social insurances for the military, government officials, teachers, workers, and farmers. As political democratization took place in the 1990s, the President, Legislators of the central government, and local leaders (mayors, council members, etc.) were all changed to direct election.[41] In order to win the support of voters, the major political parties in the campaigns usually proposed to increase social welfare expenditures, causing the government expenditures to rise. In addition, there were two compulsory social insurances implemented subsequently: National Health Insurance (NHI) in 1995 and National Pension Insurance (NPI) in 2008. Both of them have increased government spending, particularly the premium subsidies to low-income families.

Previous studies on the effects of taxes and transfer payments indicate that taxes have little income redistribution effect due to the relatively stable tax system in Taiwan. On the other hand, the large increases in social welfare and social insurance expenditures since the 1990s are more important in terms of reducing income inequality (Jao, 2000; Cheng and Lee, 2010). Therefore, we include the ratio of government social welfare and social insurance expenditures (

---

[41] For example, the 1992 legislator's election, the 1993 county mayoral election, the 1994 governor's election, and the 1996 presidential election.

$SWEXP$) to government budget (GB) ($SWEXP/GB$) as one of the influence factors of household income inequality. As shown in Figure 3.3e, $SWEXP/GB$ was about 4-7% in the 1980s and rose rapidly in the 1990s, exceeding 12% after the implementation of NHI in 1995 and increasing to around 20% after 2012.

3.4 Empirical Model

Based on the sources of household disposable income (YD), we decompose the quintile ratio (R_{YD}) to capture the influence factors of inequality and construct the empirical model to estimate their effects on R_{YD}.

According to the *RSFIE*, household disposable income includes payroll income, entrepreneurial income, property income, net transfer income, and miscellaneous receipts. We combine entrepreneurial income and property income as capital income, extract government net transfer income from net transfer income,[42] and let all other receipts as others. Therefore, household disposable income (YD) becomes the sum of payroll income (W), capital income (A), government net transfer income (GT), and others, and YD can be written as follows:

$$YD = W + A + GT + Others. \tag{1}$$

Payroll income is the main source of household income, followed by capital income.[43] Although the share of government net transfer income is relatively small, it is included as an influence factor due to its income redistribution. In the following analysis, we ignore other miscellaneous receipts since they have negligible shares in household income.

---

[42] Net transfer income includes net transfer payments from individuals, government, benefit of social insurance, enterprises, and abroad.

[43] In 2015, payroll income and capital income accounted for 56.98% and 23.38% of the total household income, respectively.

By definition, R_{YD} is the relative income of the highest 20% to the lowest 20% households. Using Eq. (1), R_{YD} can be expressed as:

$$R_{YD} = \frac{YD^H}{YD^L} = \frac{W^H + A^H + GT^H}{W^L + A^L + GT^L}, \tag{2}$$

where superscripts H and L denote the highest- and lowest-income households, respectively. R_{YD} in Eq. (2) can be decomposed into the following:

$$R_{YD} = \theta_W^L R_W + \theta_A^L R_A + \theta_{GT}^L R_{GT}, \tag{3}$$

where $R_W = W^H / W^L$, $R_A = A^H / A^L$, and $R_{GT} = GT^H / GT^L$, representing inequalities in W, A, and GT between the highest- and lowest-income households, respectively; $\theta_j^L = j^L / \left(W^L + A^L + GT^L\right)$, where $j = W, A, GT$, representing the shares of W^L, A^L and GT^L in YD^L of the lowest-income households, respectively. Using Eq. (3), we further examine the influence factors of R_W, R_A, and R_{GT} to construct our empirical model.[44]

3.4.1 Influence Factors of R_W

Household payroll income is equal to the household average wage rate (w) multiplied by the number of persons employed per household (N), i.e., $W = w \times N$. R_W can thus be expressed as:

$$R_W = W^H / W^L = w^H / w^L \times N^H / N^L, \tag{4}$$

[44] Using a similar approach as Eq. (3), Tsaur (1996) finds that increases in R_W and R_A are the main sources of household income inequality in Taiwan. However, Tsaur (1996) does not analyze the influence of government net transfer payments on inequality (R_{GT}). Changes in θ_j^L are very small in all years so the influences on inequality are negligible.

where w^H / w^L denotes the ratio of the average wage rate of the highest-income to that of the lowest-income households; N^H / N^L denotes the ratio of the number of persons employed per household of the highest-income to that of the lowest-income households. From Eq. (4), an increase in N^H / N^L or w^H / w^L will raise R_W. Figure 3.4 shows that both N^H / N^L and w^H / w^L have been greater than one and rising since 1980, causing R_W and R_{YD} to increase. Since changes in N^H / N^L are likely to be exogenous as they are due to changes in economic and social structure, changes in w^H / w^L are what needs to be explained.

$w^H / w^L > 1$ in all years implies that the average quality of labor is higher among the highest-income households. As demonstrated by Table 3.1, the highest-income households have a higher percentage of economic household heads with a bachelor's degree or higher than the lowest-income households do. Since workers with a bachelor's degree or higher are defined as skilled labor, the employed persons in the highest-income (lowest-income) households are more likely to be skilled (unskilled) labor. In general, the wage rate of skilled labor is higher than the wage rate of unskilled labor; therefore, $w^H > w^L$.

In addition, the increasing trend in w^H / w^L implies that the wage differential between skilled and unskilled workers (w_s / w_u) has increased over time. Changes in w_s / w_u depend on changes in the relative supply of and the relative demand for skilled labor. If the increase in the relative demand is larger than the increase in the relative supply, w_s / w_u will rise. We follow Johnson (1997), Chan et al. (1999), and Chen and Hsu (2001) to specify the determination of w_s / w_u.

We consider an economy in the H-O framework with two types of labor: skilled and unskilled, where both types of labor are necessary in production. There are two types of technology in production, A_1 and A_2, where A_1 is more skilled-labor-intensive than A_2. Given the level of physical capital, the technology level, and the market wage rates, firm i, where $i = 1$ and 2, will choose the two types of labor to produce output. By minimizing firm's labor costs, we can obtain the conditional demands for skilled and unskilled labor. By aggregating the economy's conditional demands, and given the aggregate supplies of skilled and unskilled labor, we can solve for the equilibrium wage rates for both types of labor. The equilibrium relative wage between skilled and unskilled labor can thus be written as:

$$w_s / w_u = f\left(N_s^s / N_u^s, A_1 / A_2, Y_1 / Y_2\right). \tag{5}$$

The above equation shows that w_s / w_u is determined by (i) the relative supply of skilled labor (N_s^s / N_u^s) and (ii) the relative demand for skilled labor, which is influenced by the relative technology level (A_1 / A_2) and the relative demand for skilled-labor-intensive output (Y_1 / Y_2).

In our empirical analysis, we use LS_C / LS_M as a proxy for N_s^s / N_u^s to represent the relative supply of skilled labor, and we use TFP as a proxy for A_1 / A_2 to reflect the domestic technological progress bias. Regarding the relative demand for output (Y_1 / Y_2), we use the ratio of net exports to GDP (NX / Y) and the ratio of employees in the service sector to those in the industrial sector (EMP_{SER} / EMP_{IND}) as proxies.[45] The influence factors of w_s / w_u will affect w^H / w^L, and thus the latter can be expressed as:

[45] The demands for domestic products consist of domestic demands and net exports (net foreign demands). Since domestic demands for traded goods are relatively stable, NX / Y can be used as a proxy for changes in the demand composition, as well as changes in the relative demand for skilled labor (Chan et al., 1999; Chen and Hsu, 2001). In addition, the increase in EMP_{SER} / EMP_{IND} indicates that the product value of the service sector grows faster than the industrial sector. Since the service sector is more skilled-labor-intensive than the industrial sector, such change in industrial structure implies an increase in the relative domestic demand for skilled-labor-intensive service products and hence in the relative demand for skilled labor (Mincer, 1993; Chan et al., 1999).

$$w^H / w^L = f\left(LS_C / LS_M, TFP, NX / Y, EMP_{SER} / EMP_{IND}\right). \tag{6}$$

Substituting Eq. (6) into Eq. (4), the possible influence factors of R_W can be written as follows:

$$R_W = R_W\left(LS_C / LS_M, TFP, NX / Y, EMP_{SER} / EMP_{IND}, N^H / N^L\right). \tag{7}$$

3.4.2 Influence Factors of R_A

The financial market development in a country not only stimulates economic growth but also improves income inequality by helping the poor to increase income through financial loans (Beck et al., 2007; Agnello et al., 2012). On the contrary, if the financial market is not well developed, a larger share of financial flows might disproportionately accrue to those with higher endowments and income. People who are already better-off are more able to invest in human/physical capital and further increase their income. As a result, financial deepening may adversely affect income inequality (IMF, 2007). Following IMF (2007), we use the ratio of $CREDIT / Y$ to represent the degree of financial development, which affects household capital income. The influence factor on R_A can be expressed as follows:

$$R_A = R_A\left(CREDIT / Y\right). \tag{8}$$

3.4.3 Influence Factors of R_{GT}

Government transfer payments lead to redistribution of household income and have a negative relationship with income inequality (Bulir, 2001; Scully, 2002). As mentioned earlier, the large increases in social welfare and social insurance expenditures are the main factors that cause income redistribution and changes in R_{GT} in Taiwan since the 1990s (Jao, 2000; Cheng and Lee, 2010). Thus, we include the ratio of government expenditures on social welfare and social insurance to government budget ($SWEXP/GB$) as the influence factor of R_{GT} :

$$R_{GT} = R_{GT}\left(SWEXP/GB\right).^{46} \tag{9}$$

3.4.4 Empirical Model and Expected Effects

Substituting Eqs. (7)-(9) into Eq. (3), we can obtain all the possible influence factors of R_{YD} . Using a linear functional form, our empirical model can be written as:

$$R_{YD} = \alpha_0 + \alpha_1 NX/Y + \beta_1 TFP + \beta_2 EMP_{SER}/EMP_{IND} + \beta_3 LS_C/LS_M$$

$$+\beta_4 N^H/N^L + \beta_5 CREDIT/Y + \beta_6 SWEXP/GB + \varepsilon_1, \tag{10}$$

where ε_1 is the error term. For simplicity, we omit time subscript t in the equation. Based on Eq. (10), we discuss the expected effects of each explanatory variable on R_{YD} as follows.

According to the S-S theorem, trade liberalization (NX/Y) improves income inequality in developing countries and deteriorates inequality in developed countries. However, the empirical results of previous studies are mixed and often inconsistent with the prediction of the S-S theorem. Moreover, since Taiwan is a MIOE, we expect an ambiguous sign for α_1 .

If technological progress (TFP) biases toward skilled labor, i.e., the relative technology level (A_1/A_2) in Eq. (5) increases, the relative wage of skilled labor and income inequality both increase. In contrast, if TFP biases toward unskilled labor, the relative wage of skilled labor and income inequality both decrease (Johnson, 1997). From Figure 3.3b, TFP has increased

[46] Both the expenditures on social welfare and total government expenditures increase over time, with the former growing faster than the latter.

substantially over time in Taiwan since 1980, so $\beta_1 > 0$ ($\beta_1 < 0$) implies that TFP biases toward skilled (unskilled) labor and raises (reduces) R_{YD}.

As shown in Figure 3.3a, the change in industrial structure (i.e., the service sector growing more rapidly than the industrial sector) results in an increase in EMP_{SER} / EMP_{IND} over time. In addition, the service sector employs more skilled labor than the industrial sector (Chan et al., 1999). These result in an increase in relative demand for skilled labor, and thus we expect a positive effect of EMP_{SER} / EMP_{IND} on R_{YD} ($\beta_2 > 0$).

Increasing prevalence in education can improve income inequality (Bourguignon, 1994; Birdsall et al., 1995; Savvides, 1998). From Figure 3.3c, the massive expansion in higher education since the 1990s in Taiwan contributes to a dramatic increase in the relative supply of skilled labor (LS_C / LS_M), which reduces the relative wage of skilled labor as well as income inequality (Chan et al., 1999; Chen and Hsu, 2001; Gindling and Sun, 2002; Vere, 2005; Lai et al., 2019). Moreover, the expansion in higher education lowers the differential in the number of persons with a bachelor's degree or higher between the highest- and lowest-income households (Table 3.1). Therefore, we expect $\beta_3 < 0$.

The greater the difference in the number of persons among households, the more unequal the household income distribution will be (Kuznets, 1981). Figure 3.4 shows that N^H / N^L has steadily increased over time, although the number of persons employed per household has gradually declined in Taiwan since 1980 (Table 3.1). We expect R_W and R_{YD} to increase and $\beta_4 > 0$.

Financial market development ($CREDIT / Y$) may improve or deteriorate income distribution between rich and poor families (Greenwood and Jovanovic, 1990; Beck et al., 2007;

IMF, 2007). From Figure 3.3d, the financial market reforms in Taiwan have led to a substantial expansion in $CREDIT/Y$ since the 1980s. If $\beta_5>0$, the result implies that high-income households get more advantage from the financial reforms (R_A and R_{YD} increase); conversely, if $\beta_5<0$, the result suggests that low-income households get more advantage (R_A and R_{YD} decrease).

Government transfer payments to households can improve income inequality through redistribution effects (Bulir, 2001; Scully, 2002). Since the 1990s, Taiwanese government has increased social welfare and social insurance expenditures, making $SWEXP/GB$ to rise substantially (Figure 3.3e). This results in significant income redistribution effects (Jao, 2000; Cheng and Lee, 2010); i.e., R_{GT} and R_{YD} decrease. Thus, we expect $\beta_6<0$.

As emphasized earlier, Taiwan is a MIOE trading with both OECD and non-OECD countries. Net exports to OECD countries and to non-OECD countries may have opposite effects on income inequality and offset each other. To account for this, we split NX/Y into NX_O/Y and NX_{NO}/Y, and modify Eq. (10) to:

$$R_{YD} = \alpha_0 + \alpha_{1\mathrm{O}} NX_{\mathrm{O}}/Y + \alpha_{1NO} NX_{NO}/Y + \beta_1 TFP + \beta_2 EMP_{SER}/EMP_{IND}$$
$$+\beta_3 LS_C/LS_M + \beta_4 N^H/N^L + \beta_5 CREDIT/Y + \beta_6 SWEXP/GB + \varepsilon_2, \quad (11)$$

where the additional subscripts O and NO of coefficient α_1 represent OECD and non-OECD countries, respectively.

According to the S-S theorem, NX_O/Y decreases R_{YD} in a less-developed country, whereas NX_{NO}/Y increases R_{YD} in a more-developed country. Coefficients α_{1O} and α_{1NO} are expected to be negative and positive, respectively. On the other hand, trading with developed countries may lead to technology transfer to a less-developed country (e.g., Lau and Wan, 1991). Learning new technology requires skilled labor, so net exports to OECD countries may increase the demand for skilled labor and raise R_{YD} in a MIOE. At the same time, commodities produced and exported to non-OECD countries require more unskilled labor. Trading with non-OECD countries may decrease R_{YD}, In this case, we expect α_{1O} to be positive and α_{1NO} to be negative. Overall, the signs of α_{1O} and α_{1NO} depend on whether the effect via comparative advantage or the effect via technology transfer dominates.

3.5 Empirical Results

3.5.1 Data, Unit Root Test, and Methodology

We use time-series data from 1980 to 2015 for our empirical analysis. Table 3.A1 in Appendix B summarizes the variables used, their definitions, data sources, and summary statistics.[47] We first conduct unit root tests with drift (τ_{μ}) on all the variables, using the ADF test (Dickey and Fuller, 1979) and the PP test (Phillips and Perron, 1988). Table 3.3 shows that all variables (except for LS_C / LS_M) contain a unit root, I(1).[48] For LS_C / LS_M, both tests indicate that it is integrated of order two, I(2). By taking the first difference, $\Delta\left(LS_C / LS_M\right)$, it becomes I(1).

If all the variables have a unit root but form a stationary linear combination, these variables are said to be cointegrated. To test for cointegration, we apply the Engle-Granger two-

[47] There are several caveats worth mentioning. First, some of the variables rely on proxies; however, these are the most commonly used measures in the literature and government reports. Second, it is likely that the high-income and low-income households are affected by the price changes differently. Using the same deflator may result in an upward bias of the effects on inequality.

[48] Both the ADF and PP tests fail to reject the null of unit root in level but reject the null in first difference.

step method (Engle and Granger, 1987), as suggested by Campbell and Perron (1991). The two steps are as follows: (i) use OLS regression to estimate the model and obtain the long-run equilibrium relationship; (ii) use the ADF test to check for stationarity of the obtained regression residuals. The null hypothesis in the second step is that the variables are not cointegrated, i.e., the residuals are not stationary. The rejection of the null implies that the variables are cointegrated. Otherwise, there is no long-run equilibrium relationship between the variables. Beyer et al. (1999) also use this method to estimate the long-run effects of trade liberalization on wage inequality in Chile.

We estimate Eqs. (10) and (11) using $\Delta\left(LS_C / LS_M\right)$ instead of LS_C / LS_M to meet the requirement of the cointegration test; that is, all variables should have the same integration order, I(1). By observing the trend in the quintile ratio (R_{YD}) in Figure 3.1, we add three dummy variables to the model: $D01$, $D09$, and $D0215$. $D01$ is equal to 1 for 2001, and 0 otherwise; $D09$ is equal to 1 for 2009, and 0 otherwise; $D0215$ is equal to 1 for 2002 to 2015, and 0 otherwise. $D01$ is used to capture the dot-com bubble in 2001, and $D09$ is used to capture the financial crisis in 2009. Both events led to economic recessions in Taiwan and increased R_{YD}. $D0215$ is used to capture the structural change in R_{YD} from 2002 to 2015. As shown in Figure 3.1, R_{YD} remained fairly stable from 2002 to 2015, unlike the previous rising trend. Such structure change is likely to be related to the industrial structure. As shown in Figures 3.3a and 3.4, the distribution of the number of employees in the labor market (EMP_{SER} / EMP_{IND}) and the ratio of household average wage rate (w^H / w^L) have both remained stable since 2002.

3.5.2 Estimation Results

Table 3.4 reports the estimation results of Eqs. (10) and (11), where the last row shows the outcomes of the cointegration tests on the residuals ($\hat{\varepsilon}_1$, $\hat{\varepsilon}_2$). The ADF statistics (τ_μ) in these two equations reject the null of no cointegration, indicating that R_{YD} and all the explanatory variables are cointegrated. Hence, the estimated coefficients represent the long-run effects of the explanatory variables on R_{YD}. The coefficients of $D01$ and $D09$ are both positive and statistically significant. This suggests that the employed persons of the lowest-income households (unskilled labor) were more likely to lose jobs in the 2001 and 2009 recessions, causing R_{YD} to rise. The dummy variable $D0215$ has a positive and statistically significant coefficient, which implies that R_{YD} does not change much since the industrial structure remains stable during this time period.

Considering the effects of trade liberalization on income inequality, we show that NX/Y has a positive and statistically significant effect on R_{YD}, which is consistent with the findings in Chan et al. (1999) and Chen and Hsu (2001).[49] When distinguishing the effects between NX_O/Y and NX_{NO}/Y on R_{YD}, we find that NX_O/Y has a positive and statistically significant effect, whereas the effect of NX_{NO}/Y is negative and statistically insignificant.[50] Overall, our results show that trade liberalization increases income inequality in the long run, and it is mainly attributed to net exports to OECD countries. As a MIOE, Taiwan is less skilled-labor-intensive than most OECD countries and more skilled-labor-intensive than most non-OECD countries. Based on the S-S theorem, net exports to OECD countries decrease the relative wage of skilled labor as well as income inequality. On the other hand, net exports to non-OECD countries

[49] They find a positive effect of NX/Y on the relative wage of skilled labor in Taiwan.
[50] The signs of the effects of NX_O/Y and NX_{NO}/Y on R_{YD} are the same as those obtained by Chen and Hsu (2001), which uses the relative wage of skilled labor as the outcome variable.

increase the relative wage of skilled labor and deteriorate income inequality. Therefore, our results are different from the prediction using the S-S theorem.

Why are the trade effects of NX_O / Y and NX_{NO} / Y inconsistent with the S-S theorem? As Chen and Hsu (2001) point out, the products exported to OECD countries are among the most complicated varieties in Taiwan. Net exports to OECD countries could increase relative demand for skilled labor for two reasons. First, through learning-by-doing, trading with more-developed countries helps upgrade the technology level and thus increases the demand for skilled labor (e.g., Chuang, 1998). Second, according to the capital accumulation-outsourcing hypothesis (Feenstra and Hanson, 1996, 1997), developing countries upgrade the range of intermediate inputs that they produce and export through outsourcing of multinationals, which increases their demand for skilled labor. While more skilled workers are employed in the sectors that export to OECD countries, there are fewer skilled workers available to the sectors that export to non-OECD countries. In fact, exports to non-OECD countries (e.g., China) are usually more unskilled-labor-intensive products in Taiwan, thereby demanding more unskilled labor.

Unfortunately, there is no data available for classifying the composition of exports to and imports from OECD and non-OECD countries in Taiwan. To get a sense of the composition, we use data from three major trade partners of Taiwan in 1998 and 2010. These include two OECD countries (the U.S. and Japan) and one non-OECD country (China). As shown in Table 3.5, these three countries together accounted for more than 45% of total exports and imports in Taiwan. In both years, the shares of skilled-labor-intensive products in exports to the U.S. (73.7% and 77.9% in 1998 and 2010, respectively) and Japan (62.3% and 73.6%) were higher than those to China (39.0% and 59.6%). Similarly, the shares of skilled-labor-intensive products in imports from the U.S. (66.0% and 62.5%) and Japan (78.9% and 76.8%) were much higher than those from China (9.0% and 19.0%). Among the items of exports and imports, machinery and electrical equipment are predominant. In both years, exports of machinery and electrical equipment as a share of exports to the U.S. (58.7% and 55.7%) and Japan (46.4% and 53.8%) were much higher than those to China (10.0% and 38.7%). More noticeably, the shares of machinery and electrical equipment in imports were 40.8% and 39.3% from the U.S. and 51.6% and 43.8% from Japan; however, none of them were imported from China. These findings suggest that Taiwan imported skilled-labor-intensive machinery and electrical equipment from technologically advanced countries. They were used to produce skill-labor-intensive products and were mainly exported to OECD countries. This skill composition is consistent with the learning model (e.g., Pissarides, 1997) and tends to increase the relative wage of skilled workers (Chen and Hsu, 2001).

In contrast, Taiwan seemed to import unskilled-labor-intensive products from developing countries and export more unskilled-labor-intensive products to non-OECD countries. Indeed, the increase in net exports to non-OECD countries since the 1990s were largely attributed to growing trade with China. The large foreign direct investment (FDI) in China in the past decades was mainly due to cheaper unskilled labor in China, and most technology was less-advanced. Accompanied with FDI, the Taiwanese multinationals also imported less-advanced intermediates from Taiwan to produce products for sales in China. As a result, an increase in net exports to China raised the relative demand for unskilled labor in Taiwan and decreased income inequality, which is in line with the finding from non-OECD countries.

In terms of the effects of other explanatory variables on R_{YD}, the sign and statistical significance of the coefficients are the same between Eqs. (10) and (11), and their magnitudes are close. This suggests that the estimated effects of these variables are robust. Technological progress (TFP) has a positive and statistically significant effect on R_{YD} (β_1 is around 0.04), which is consistent with Chan et al. (1999) and Lai et al. (2019). IMF (2007) argues that technological progress biased toward skilled labor is the most important cause of the deterioration of income inequality in both developed and developing countries. Our results also suggest that changes in industrial development policy since 1980, which leads to technological progress biased toward skilled labor, is an important factor of R_{YD}.

The effect of EMP_{SER} / EMP_{IND} on R_{YD} is positive and statistically significant (β_2 is around 1.5), as expected. This indicates that faster growth in the service sector relative to the industrial sector since the late 1980s absorbs more employees and in turn increases the relative demand and the relative wage of skilled labor (Chan et al., 1999),[51] resulting in an increase in R_{YD}. The massive expansion in higher education since the 1990s increases the relative supply of skilled labor. We find that $\Delta\left(LS_C / LS_M\right)$ has a negative effect on R_{YD}, but it is statistically insignificant.[52] The effect of N^H / N^L on R_{YD} is negative but statistically insignificant.[53] As demonstrated in Figure 3.1, the quintile ratio (R_{YD}) remained relatively stable after 2002. However, Figure 3.4 shows that N^H / N^L kept increasing while w^H / w^L became stable and then declined after 2014. These trends imply that the main influence factor on inequality in household payroll income (R_W) is w^H / w^L but not N^H / N^L.

The financial market development ($CREDIT / Y$) has a positive and statistically

[51] Mincer (1993) also finds that the rising trend in service employment reflects more demand for skilled labor, which can partly explain the educational wage differential in the US.
[52] This is consistent with the finding in Chen and Hsu (2001), which shows a statistically insignificant negative effect on the relative wage of skilled labor, as well as the finding in Lai et al. (2019), which shows a negative effect on income inequality.
[53] Lai et al. (2019) also find an statistically insignificant effect of N^H / N^L on R_{YD}.

significant effect on R_{YD} (β_5 is around 0.6). With the implementation of several financial reforms in Taiwan since the early 1980s, our results show that the high-income households seem to gain more advantage from those reforms, and adversely affect income inequality. Finally, we find that $SWEXP/GB$ has a negative and statistically significant effect on R_{YD} (β_6 is about -3.7), which is consistent with previous studies and the expectation. This result suggests that the large increases in social welfare and social insurance expenditures indeed lead to income redistribution effects and reduce income inequality.

There are two key takeaways based on the results in Table 3.4. First, trade liberalization overall raises income inequality, which is mainly attributed to net exports to OECD countries. Second, among the domestic reforms, technological progress in favor of skilled labor, change in industrial structure, and financial market reforms, all increase inequality. Although the massive expansion in higher education increases the relative supply of skilled labor, it does not have a significant effect on inequality. Increases in social welfare and social insurance expenditures help reduce inequality.

3.5.3 Robustness Analysis

We assess the robustness of the results by estimating the following alternative specifications. First, we separately analyze the effects of exports and imports to ensure the robustness of the combined effects of net exports on R_{YD}. The results are shown in Columns (1) and (2) of Table 3.6. In Column (1), we find a positive effect of total exports (X/Y) and a negative effect of total imports (M/Y), which is consistent with the previous result that the effect of NX/Y on R_{YD} is positive. In Column (2), we further estimate the effects of exports and imports with OECD countries (X_O/Y, M_O/Y) and those with non-OECD countries (X_{NO}/Y,

M_{NO}/Y). As shown, the effect of X_O/Y is positive and statistically significant, whereas the effect of M_O/Y is negative but statistically insignificant. These two effects coincide with the skill composition of exports to and imports from OECD countries in Table 3.5. Since Taiwan exports skilled-labor-intensive products to OECD countries and imports skilled-labor-intensive products from OECD countries, R_{YD} would be increased by X_O/Y but decreased by M_O/Y. Altogether, the effects of X_O/Y and M_O/Y support the positive and statistically significant effect of NX_O/Y on R_{YD} in Table 3.4. On the other hand, both exports and imports with non-OECD countries (X_{NO}/Y and M_{NO}/Y) have negative effects on R_{YD}, but they are statistically insignificant. Since the magnitude of the coefficient of X_{NO}/Y (0.9836) is a greater than the coefficient of M_{NO}/Y (0.3416) in absolute terms, this can explain why the combined effect of NX_{NO}/Y is negative but statistically insignificant in Table 3.4. Overall, the separate effects of exports and imports in Table 3.6 are consistent with the effects of net exports in Table 4. In addition, the results of all the other variables are robust.

Second, we use the Gini coefficient (G_{YD}) as an alternative indicator of income inequality and re-estimate Eqs. (10) and (11). The results are shown in Columns (3) and (4) of Table 3.6.[54] Since G_{YD} has a value between 0 and 1 (where 0 denotes absolute equality, and 1 denotes absolute inequality), the magnitudes of the estimated coefficients would be different from those in Table 3.4. Regarding the effects of trend, both NX/Y and NX_O/Y have a positive and statistically significant effect, whereas NX_{NO}/Y has a positive but statistically insignificant

[54] We only report the estimated coefficients with their signs and statistical significance to save space. The detailed results are available upon request.

effect. These trade effects are consistent with the finding in Table 3.4, which suggests that trade liberalization overall increases inequality and that the effect is mainly attributed to net exports to OECD countries. With respect to domestic reforms, TFP and EMP_{SER} / EMP_{IND} have a positive and statistically significant effect, and $SWEXP / GB$ has a negative and statistically significant effect. The effect of $\Delta(LS_C / LS_M)$ is still insignificant, although its sign changes from negative to positive. The effect of $CREDIT / Y$ is still positive but becomes statistically insignificant, suggesting that the financial reforms may not be as influential as other reforms. In general, the results using G_{YD} as an indicator of inequality are consistent with the results in Table 3.4.

3.5.4 Contributions to Income Inequality

Since the explanatory variables are measured in different units (ratio, index, percentage, etc.), we cannot directly assess the contributions of each variable to income inequality from the estimated coefficients in Table 3.4. To address this issue, we use the results in Table 3.4 (the marginal long-run effect on R_{YD}) and the mean values in Table 3.A1 to calculate the long-run mean effect of the following six variables that have a statistically significant effect on R_{YD}: NX / Y, NX_O / Y, TFP, EMP_{SER} / EMP_{IND}, $CREDIT / Y$ and $SWEXP / GB$. With these mean effects, we can evaluate the relative importance of trade liberalization and domestic reforms on R_{YD} during the sample period.

As shown in Table 3.7, the total net exports have a positive long-run effect on R_{YD}. A one percentage point increase in NX / Y leads to a 1.4516 percent increase in R_{YD}, *ceteris paribus*. However, its long-run mean impact on R_{YD} was only 0.0994. Noticeably, NX_O / Y had a positive long-run effect on R_{YD} (1.7321), but its long-run mean impact was negative (-0.0113). This is because Taiwan has trade deficits with OECD countries since 1990 (with a mean value of -0.65%), which reversely leads to a decrease in inequality.

Technological progress has the largest long-run mean effect on R_{YD}, which is about 3.73. EMP_{SER} / EMP_{IND} has the second largest long-run mean effect, and $CREDIT / Y$ also contributes to the rising inequality with a long-run mean effect of 0.71. Among domestic reforms, only $SWEXP / GB$ decreases income inequality, but it is less influential compared to the other domestic reforms that increase R_{YD}.

Based on the results in Tables 3.4 and 3.7, we can conclude that trade liberalization overall raises household income inequality in Taiwan from 1980 to 2015, and this adverse effect is mainly due to net exports to OECD countries. However, since trade liberalization leads to trade diversion, causing trade deficits with OECD countries that in turn reducing inequality, the overall impact of trade liberalization on inequality is relatively small. Considering all the influence factors of income inequality, we show that domestic reforms, particularly technological progress in favor of skilled labor, industrial structure change, and financial reforms, are the main influence factors of the rising inequality.

3.6 Concluding Remarks

This study estimates the effects of trade liberalization and domestic reforms on income inequality in Taiwan from 1980 to 2015. We construct an empirical model by decomposing the sources of household income in the quintile ratio. Since there are many MIOEs that simultaneously trade with more- and less-developed countries, it is important to distinguish the trade effects between these two types of trade partners. We investigate the case of Taiwan and separately estimate of the effects of net exports to OECD and non-OECD countries on income inequality. We also examine the contributions of various influence factors to the rising inequality.

Our results suggest that trade liberalization overall increases inequality, but the long-run mean impact is relatively small. Domestic reforms, such as technological progress in favor of skilled labor, change in industrial structure, and financial reforms, have negative and larger impacts on inequality than trade liberalization, indicating that they might be the main influence factors of inequality.

Our findings on the trade effects of OECD and non-OECD countries diverge from the prediction of the Stolper-Samuelson theorem. We show that the negative effects of trade liberalization on inequality are mainly resulted from trade with OECD countries. Although net exports to OECD countries increase inequality, trade diversion due to liberalization leads to trade deficits with OECD countries and in turn decreases the overall impact on inequality in the long run. Our paper highlights the importance of distinguishing between trade partners, which might be one possible reason for the mixed results in the literature.

Similar to Taiwan, many developing countries have endeavored in capital accumulation and technological progress in order to promote economic growth and increase income levels. Accompanied with economic growth, industrial structure has also changed over time, in particular the rapid growth of the service sector. We demonstrate that both technological progress in favor of skilled labor and the growth of the service sector increase the relative demand for skilled labor, and they substantially increase income inequality. Although the government increases social welfare and social insurance transfer payments to low-income families, their effects on inequality are limited. In the process of economic development, our study shows that economic growth and income equality seem to be an inevitable trade-off.

References

Agnello, L., Mallick, S. K., Sousa, R. M., 2012. Financial Reforms and Income Inequality. Econ. Lett. 116, 583-587.

Asteriou, D., Dimelis, S., Moudatsou, A., 2014. Globalization and Income Inequality: A Panel Data Econometric Approach for EU27 Countries. Econ. Modelling. 36, 592-599.

Barro, R. J., 2000. Inequality and Growth in a Panel of Countries. J. Econ. Growth 5(1), 5-32.

Beck, T., Demirguc-Kunt, A., Levin, R., 2007. Finance, Inequality and the Poor. J. Econ. Growth 12, 27-49.

Behrman, J. R., Birdsall, N., Szekely, M., 2003, updated 2006. Economic Policy Changes and Wage Differential in Latin America. Center for Global Dev. Working Paper 29, Washington, D.C.

Berggren, N., 1999. Economic Freedom and Equality: Friends or Foes? Publ. Choice 100, 203-223.

Bergh, A., Nilsson, T., 2010. Do Liberalization and Globalization Increase Income Inequality? Eur. J. Polit. Econ. 26, 488-505.

Berman, E., Bound, J., Griliches, Z., 1994. Changes in the Demand for Skilled Labor Within US Manufacturing: Evidence from Annual Survey of Manufacturers. Q. J. Econ. 109, 367-398.

Beyer, H., Rojas, P., Vergara, R., 1999. Trade Liberalization and Wage Inequality. J. Dev. Econ. 59(1), 103-123.

Birdsall, N., Ross, D., Sabot, R.-H., 1995. Inequality and Growth Reconsidered: Lessons from East Asia. World Bank Econ. Rev. 9(3), 477-508.

Bourguignon, F., 1994. Growth, Distribution, and Human Resources, in: Ranis, G. ed., En Route to Modern Economic Growth, 43-70. Johns Hopkins University Press.

Bourguignon, F., Fournier, M., Gurgand, M., 2001. Fast Development with a Stable Income Distribution: Taiwan, 1979-94. Rev. Income Wealth 47(2), 139-163.

Bulir, A., 2001. Income Inequality: Does Inflation Matter? IMF Staff Pap. 48(1), 139-159.

Campbell, J., Perron, P., 1991. Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots. NBER Macroecon. Ann. 1991, vol. 6. The MIT Press, 141-201.

Carter, J. R., 2006. An Empirical Note on Economic Freedom and Income Inequality. Publ. Choice 130, 163-177.

Chan, V.-L., Chen, L.-T., Hu, S.-C., 1999. Implications of Technology and Education for Wage Dispersion: Evidence from Taiwan. In: G. Ranis, S-C Hu and Y-P Chu (eds.) The

Economics and Political Economy of Taiwan's Development into the 21st Century. Aldershot: Edward Elger. 247-284.

Chen, B. L., Hsu, M, 2001. Time-Series Wage Differential in Taiwan: The Role of International Trade. Rev. Dev. Econ. 5(2), 336-354.

Cheng, P. C. R., Lee, Y., 2010. Global and Local Analysis of the Redistributive Effect of Taxes and Benefits: The Case of Taiwan, 1976-2006. Taiwan Econ. Rev. 38(2), 233-288. (in Chinese)

Chuang, Y.-C., 1998. Learning by Doing, the Technology Gap, and Growth. Int. Econ. Rev. 39(3), 697-721.

Dickey, D. A., Fuller, W. A., 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Am. Stat. Assoc. 74, 427-431.

Dreher, A., Gaston, N., 2008. Has Globalization Increased Inequality? Rev. Int. Econ. 16(3), 1-21.

Engle, R. F., Granger, C. W. J., 1987. Co-integration and Error Correction: Represention, Estimation, and Testing. Econometrica 55(2), 251-276.

Feenstra, R. C., Hanson, G. H., 1996. Globalization, Outsourcing, and Wage Inequality. Am. Econ. Rev. 86(2), 240-245.

Feenstra, R. C., Hanson, G. H., 1997. Foreign Direct Investment and Relative Wages: Evidence from Mexico's Maquiladoras. J. Int. Econ. 42(3), 371-393.

Galiani, S., Sanguinetti, P., 2003. The Impact of Trade Liberalization on Wage Inequality: Evidence from Argentina. J. Dev. Econ. 72(2), 497-513.

Gindling, T. H., Sun, W., 2002. Higher Education Planning and the Wages of Workers with Higher Education in Taiwan. Econ. Educ. Rev. 21, 153-169.

Goldberg, P. K., Pavcnik, N., 2007. Distribution Effects of Globalization in Developing Countries. J. Econ. Lit. 45, 39-82.

Greenwood, J., Jovanovic, B., 1990. Financial Development, Growth, and the Distribution of Income. J. Polit. Econ. 98(5), 1076-1107.

Hanson, G., Harrison, A., 1999. Trade Liberalization and Wage Inequality in Mexico. Ind. Labor Relat. Rev. 52(2), 271-288.

Herault, N., 2007. Trade Liberalization, Poverty and Inequality in South Africa: A Computable General Equilibrium-Micro Simulation Analysis. Econ. Record 83(2), 317-328.

IMF, 2007. World Economic Outlook-Globalization and Inequality, Chapter 4, October, Washington D.C.

Jao, C.-C., 2000. The Impact of Tax Revenue and Social Welfare Expenditure on Income Distribution in Taiwan. J. Asia Pacific Econ. 5(1/2), 73-90.

Johnson, G. E., 1997. Changes in Earnings Inequality: The Role of Demand Shifts. J. Econ. Perspect. 11, 41-54.

Kumar, U., Mishra, P., 2008. Trade Liberalization and Wage Inequality: Evidence from India. Rev. Dev. Econ. 12(2), 291-311.

Kuznets, S., 1981. Size of Households and Income Disparities. In: Simon, J. L., Lindert, P. H. (Eds.), Research in Population Economics, vol. 3, 1-40, JAI Press.

Lai, M.-Y., Liu, J., Liu, Z.-S. 2019. Economic Globalization and Household Income Inequality: An Empirical Study of Taiwan (in Chinese). Taiwan Econ. Forecast Policy. 49(2), 1-46.

Lawrence R. Z., Slaughter, M. J., 1993. International Trade and American Wages in the 1980s: Giant Sucking Sound or Small Hiccup. Brookings Pap. Econ. Act., 161-210.

Lau, M.-l., Wan, H., 1991. The Theory of Growth and Technology Transfer: Experience from the East Asian Economies. Seoul J. Econ. 4, 109-122.

Mah, J. S., 2003. A Note on Globalization and Income Inequality-The Case of Korea, 1975-1995. J. Asian Econ. 14(1), 157-164.

Mcnabb, R., Said, R., 2013. Trade Openness and Wage Inequality: Evidence for Malaysia. J. Dev. Stud. 49(8), 1118-1132.

Milanovic, B., 2005. Can We Discern the Effect of Globalization on Income Distribution? Evidence from Household Surveys. World Econ. Rev. 19, 21-44.

Mincer, J., 1993. Human Capital, Technology, and the Wage Structure: What Do Time Series Show? In: Studies in Human Capital. Edward Elgar, 366-405.

National Development Council, Republic of China (Taiwan), 2016. Economic Development, R.O.C. (Taiwan).

Phillips, P. C. B., Ouliaris, S., 1990. Asymptotic Properties of Residual Based Tests for Cointegration. Econometrica 58(1), 165-193.

Phillips, P. C. B., Perron, P., 1988. Testing for a Unit Root in Time Series Regression. Biometrika 75(2), 335-346.

Pissarides, C. A., 1997. Learning by Trading and the Returns to Human Capital in Developing Countries. World Bank Econ. Rev. 11(1), 17-32.

Reuveny, R., Li, Q., 2003. Economic Openness, Democracy, and Income Inequality: An Empirical Analysis. Comp. Polit. Stud. 36(5), 575-601.

Sato, S., Fukushige, M., 2009. Globalization and Income Inequality in the Short and Long Run: The Case of South Korea 1975-1995. J. Asian Econ. 20(1), 62-68.

Savvides, A., 1998. Trade Policy and Income Inequality: New Evidence. Econ. Lett. 61(3), 365-372.

Scully, G. W., 2002. Economic Freedom, Government Policy and the Trade-off between Equality and Growth. Publ. Choice 113, 77-96.

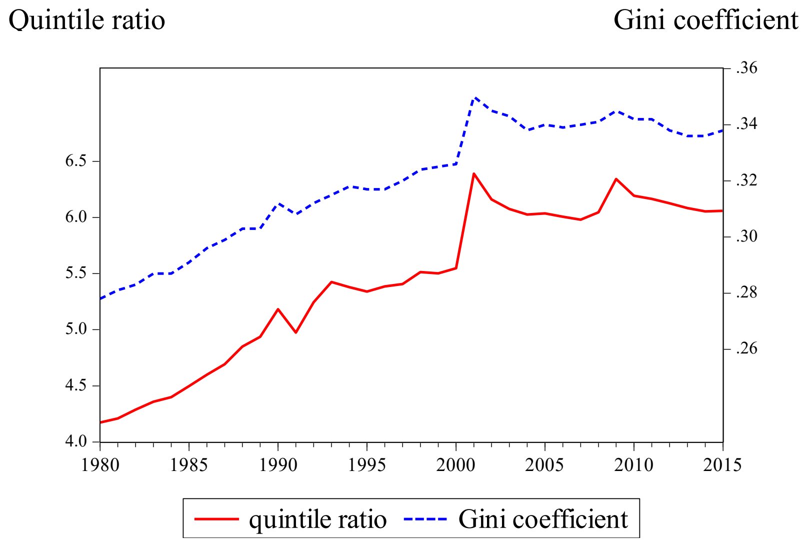
Stolper, W., Samuelson, P., 1941. Protection and Real Wages. Rev. Econ. Stud. 9(1), 58-73.

Tsaur, T.-W., 1996. An Analysis of the Variation of Family Income Inequality in Taiwan, 1980-1993. J. Soc. Sci. Phil. 8(2), 181-219. (in Chinese)

Vere, J. P., 2005. Education, Development, and Wage Inequality: The Case of Taiwan. Econ. Dev. Cult. Change 53(3), 711-735.

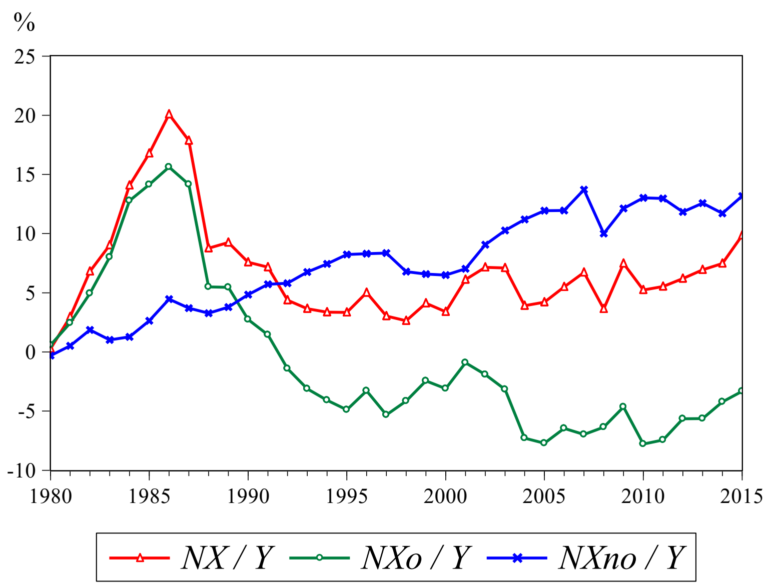
Wood, A., 1997. Openness and Wage Inequality in Developing Countries: The Latin American Challenge to East Asia Conventional Wisdom. World Bank Econ. Rev. 11(1), 33-57.

Figure 3.1: Quintile Ratio and Gini Coefficient



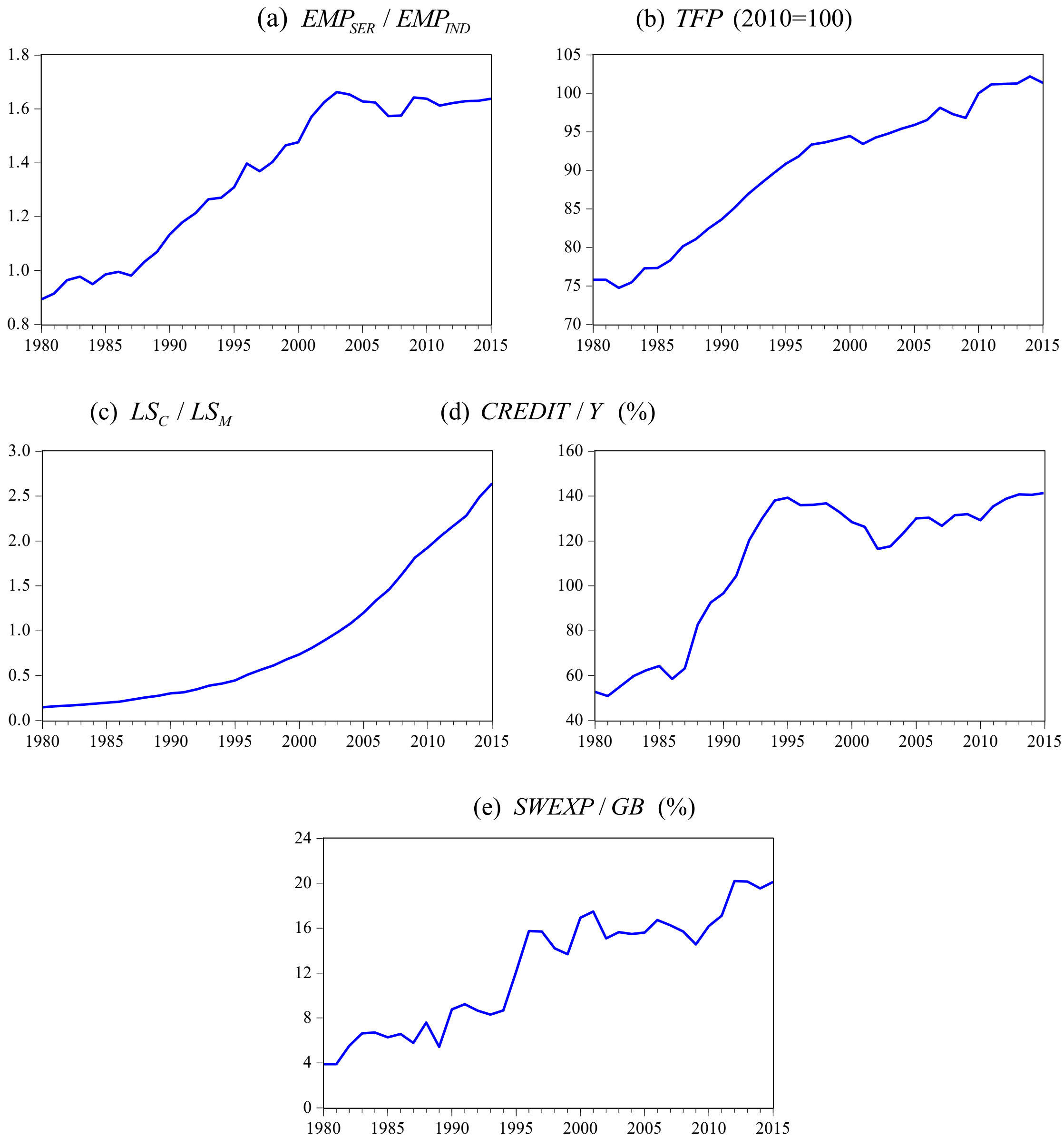
Source: *RSFIE*, DGBAS, ROC (Taiwan).

Figure 3.2: Ratios of Net Exports to GDP



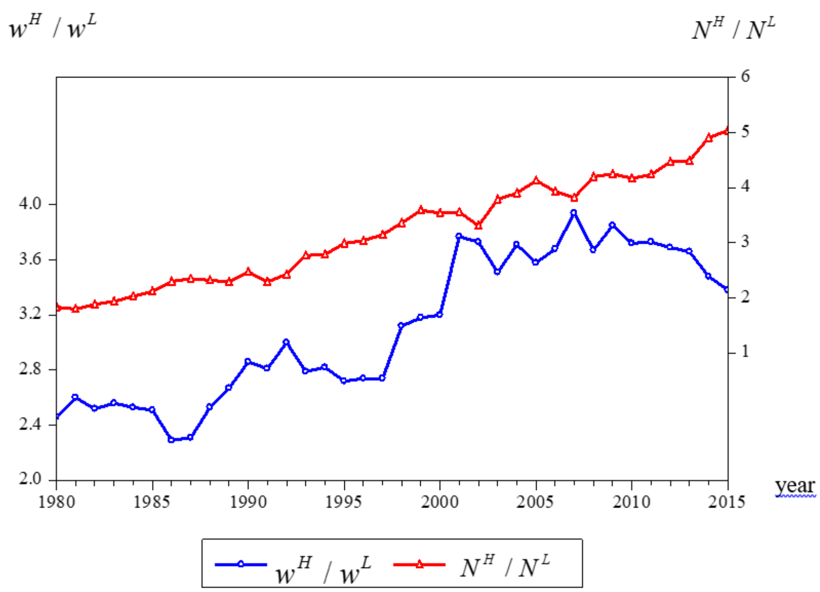
Note: NX / Y denotes the ratio of total net exports to GDP; NX_O / Y denotes the ratio of net exports to OECD countries to GDP; NX_{NO} / Y denotes the ratio of net exports to non-OECD countries to GDP.
Source: *AREMOS Databank*, Taiwan Economic Data Center.

Figure 3.3: Trends of the Related Variables of Domestic Reforms



Note: The horizontal axis in all panels denotes the calendar year (1980-2015). The vertical axis in each panel denotes the corresponding variable shown. EMP_{SER} / EMP_{IND} denotes the ratio of the number of employees in the service sector to the number of employees in the industrial sector. TFP denotes the total factor productivity index in industrial and service sectors. LS_C / LS_M denotes the ratio of the number of employees with a bachelor's degree or higher to those with mandatory education or below. $CREDIT / Y$ denotes the ratio of credit (loans and discounts) to the private sector by monetary institutions to GDP. $SWEXP / GB$ denotes the ratio of social welfare and social insurance expenditures to government budget. More details about these variables are in Table A1 in Appendix B.

Figure 3.4: Ratio of the Number of Persons Employed and Average Wage Rate Differential between Highest- and Lowest-Income Households



Note: N^H / N^L denotes the ratio of the number of persons employed, and w^H / w^L denotes the average wage rate differential. Household average wage rate (w) = household payroll income (W)/the number of persons employed per household (N).
Source: Household payroll income and the number of persons employed per household are from the *RSFIE*, DGBAS.

Table 3.1: Average Number of Persons Employed Per Household and Household Head Educational Attainment of First and Fifth Quintiles

Household characteristics	1st quintile	5th quintile	Ratio
Number of persons per household (person)			
1980	3.62	5.80	1.60
1990	2.70	5.09	1.89
2000	1.99	4.65	2.34
2015	1.71	4.21	2.46
Number of persons employed per household (person)			
1980	1.46	2.65	1.82
1990	1.06	2.59	2.44
2000	0.68	2.41	3.54
2015	0.46	2.32	5.04
Percentage of economic household heads with a bachelor's degree or higher (%)			
1980	1.23	19.93	16.20
1990	1.81	25.53	14.10
2000	2.63	30.42	11.57
2015	5.72	45.14	7.89

Source: *RSFIE*, DGBAS, ROC (Taiwan).

Table 3.2: Summary of Trade Policy Changes and Domestic Reforms in Taiwan

Policies	1980s	1990s	2000s~
Trade liberalization:			
	1984: adopt trade liberalization policy 1987: relax foreign exchange control 1989: implement managed floating exchange rate system	1993: allow indirect trade with China	2002: allow direct trade with China 2002: become a member of WTO
Domestic reforms:			
Industrial development	1960~1990: Statute for the Encouragement of Investment (SEI)	1991: Statute for Upgrading Industry (SUI)	2010: Statute for Industrial Innovation (SII)
Higher education	1985: allow establishment of new universities/colleges	1990s: expand the number of universities and college students	2000s: further increase the number of universities and college students
Financial market	1984: allow domestic banks to set more branches and foreign banks to set branches in Taiwan 1989: liberalize interest rates of banks	1991: approve the set-up of 15 new banks	2002: first financial reform 2004: second financial reform
Social welfare and social insurance		1990s: increase spending on social welfare 1995: implement National Health Insurance (NHI)	2008: implement National Pension Insurance (NPI)

Table 3.3: Unit Root Tests

Variable	ADF test		PP test	
	Level	1st difference	Level	1st difference
R_{YD}	-1.6392(0)	-7.0157***(0)	-1.9887(8)	-7.5083***(8)
NX / Y	-2.1147(0)	-4.7989***(0)	-2.5119(8)	-4.7166***(8)
NX_O / Y	-1.0175(0)	-4.3307***(0)	-1.3046(8)	-4.2162***(8)
NX_{NO} / Y	-1.3752(0)	-7.0371***(0)	-1.4433(8)	-8.5193***(8)
TFP	-1.0947(0)	-4.9315***(0)	-1.0742(8)	-4.9808***(8)
N^H / N^L	0.3109(0)	-7.1021***(0)	1.3769(8)	-8.5317***(8)
EMP_{SER} / EMP_{IND}	-1.5435(0)	-4.8385***(0)	-1.4474(8)	-4.9557***(8)
LS_C / LS_M	-0.5569(5)	0.0365(4)	8.0075(8)	-0.7059(8)
$\Delta\left(LS_C / LS_M\right)$	0.2359(4)	-4.7517***(3)	-0.5676(8)	-10.7930***(8)
$CREDIT / Y$	-2.0804(1)	-3.4408**(0)	-1.6347(8)	-3.4988**(8)
$SWEXP / GB$	-1.0157(0)	-5.1741***(2)	-0.8279(8)	-6.4625***(8)

Note: Figures in the parentheses following test statistics τ_μ denote lag periods (p); for ADF test, p is chosen by Schwarz Information Criterion (SIC) (max p =8); for PP test, p = 8. Critical values of 1%, 5%, and 10% significant levels for ADF test are -3.64, -2.96, and -2.62, respectively; those for PP test are -3.65, -2.95, and -2.62, respectively. ***, **, and * denote significance in 1%, 5%, and 10% levels, respectively.

Table 3.4: Cointegration Tests and Estimations

Dependent Variable: R_{YD}				
	Eq. (10)		Eq. (11)	
Explanatory Variables	coefficient	s.e.	coefficient	s.e.
Const.	0.1982	0.9136	-0.1734	0.9686
NX / Y	1.4516***	0.4420		
NX_O / Y			1.7321***	0.5069
NX_{NO} / Y			-0.3284	1.6587
TFP	0.0366**	0.0148	0.0414**	0.0153
EMP_{SER} / EMP_{IND}	1.5227***	0.3270	1.4805***	0.3277
$\Delta(LS_C / LS_M)$	-0.2276	1.0060	-0.4065	1.0141
N^H / N^L	-0.1382	0.0985	-0.1505	0.0986
$CREDIT / Y$	0.5052**	0.2283	0.6348**	0.2553
$SWEXP / GB$	-3.6800***	1.2234	-3.6915***	1.2177
$D01$	0.8092***	0.1170	0.8188***	0.1168
$D09$	0.1943*	0.1126	0.2147*	0.1135
$D0215$	0.2949**	0.1161	0.3922**	0.1449
Adj. R^2	0.9846		0.9848	
Durbin-Watson stat.	2.1953		2.2667	
Obs.	36		36	
Residual ADF stat. (τ_μ)	-6.3678***(0)		-6.6153***(0)	

Note: Figures in the parentheses following residual ADF test statistics (τ_μ) denote lag periods (p) and are chosen by SIC (max p =8); critical values of 1%, 5%, and 10% significant levels are -5.2812, -4.7101, and -4.4309, respectively (Phillips and Quliaris, 1990). ***, **, and * denote significance in 1%, 5%, and 10% levels, respectively.

Table 3.5: Composition of Exports and Imports of Major Trade Partners (Unit: %)

	1998			2010		
	U.S.	Japan	China	U.S.	Japan	China
Share in Taiwan's total exports	26.6	8.4	17.9	11.5	6.6	28.0
Share in Taiwan's total imports	18.0	25.8	3.9	10.1	20.7	14.3
Share of skilled-labor-intensive products in exports	73.7	62.3	39.0	77.9	73.6	59.6
Share of skilled-labor-intensive products in imports	66.0	78.9	9.0	62.5	76.8	19.0
Share of machinery and electrical equipment in exports	58.7	46.4	10.0	55.7	53.8	38.7
Share of machinery and electrical equipment in imports	40.8	51.6	0.0	39.3	43.8	0.0

Note: Following Chen and Hsu (2001), we classify machinery and electrical equipment, chemicals, transportation equipment, and basic metals and articles thereof as skilled-labor-intensive products.
Source: The shares of Taiwan's total exports and imports in 1998 are from *Taiwan Statistical Data Book* (for the US and Japan), *Cross-Strait Economic Statistics Monthly* (for China), and Chen and Hsu (2001). Data in 2010 are from *Taiwan Statistical Data Book* (2016).

Table 3.6: Robustness Tests

Dependent Variable	R_{YD}		G_{YD}	
Explanatory Variables	(1)	(2)	(3)	(4)
Const.	0.0902	-0.3185	0.1304***	0.1229***
X / Y	1.1324§			
M / Y	-1.4753***			
X_O / Y		1.5704**		
M_O / Y		-0.2308		
X_{NO} / Y		-0.9836		
M_{NO} / Y		-0.3416		
NX / Y			0.0636***	
NX_O / Y				0.0692***
NX_{NO} / Y				0.0276
TFP	0.0425**	0.0460***	0.0013**	0.0014***
EMP_{SER} / EMP_{IND}	1.3798***	0.9417**	0.0738***	0.0729***
$\Delta(LS_C / LS_M)$	-0.1256	-0.1722	0.0157	0.0121
N^H / N^L	-0.1389	-0.1274	-0.0074*	-0.0076*
$CREDIT / Y$	0.4160†	0.7072**	0.0057	0.0083
$SWEXP / GB$	-3.5528***	-2.7806**	-0.1348***	-0.1351***
$D01$	0.8207***	0.9257***	0.0200***	0.0202***
$D09$	0.1895§	0.2662**	0.0010	0.0014
$D0215$	0.3340**	0.6008***	0.0025	0.0044
Adj. R^2	0.9842	0.9865	0.9838	0.9840
Durbin-Watson stat.	2.2140	2.4038	1.9816	1.9798
Obs.	36	36	36	36
Residual ADF stat. (τ_μ)	-6.4548***(0)	-7.1956***(0)	-5.7556***(0)	-5.7606***(0)

Notes: See Table 4. § and † represent 11% and 14% significance levels, respectively. Note that all the newly added variables (X/Y, M/Y, X_O/Y, M_O/Y, X_{NO}/Y, M_{NO}/Y, G_{YD}) have a unit root, I(1), which meets the requirement of a cointegration test. As shown in the last row, all the residual ADF tests reject the null of no cointegration, suggesting that R_{YD} and G_{YD} both have a long-run equilibrium relationship with their explanatory variables.

Table 3.7: Long-Run Mean Impacts on Income Inequality

Explanatory variable	Marginal effect	Mean	Long-run mean effect
NX / Y	1.4516	0.0685	0.0994
NX_O / Y	1.7321	-0.0065	-0.0113
TFP	0.0414	89.9973	3.7259
EMP_{SER} / EMP_{IND}	1.4805	1.3491	1.9973
$CREDIT / Y$	0.6348	1.1117	0.7057
$SWEXP / GB$	-3.6915	0.1239	-0.4574

Note: The long-run marginal effects are from Table 4, and the mean values are from Table A1.

Appendix

Table 3.A1: Variable Definition, Source, and Summary Statistics

Variable	Definition	Source	Mean	S.D.
R_{YD}	Quintile ratio, i.e., ratio of disposable income share of the highest 20% to that of the lowest 20% households	*Report on the Survey of Family Income and Expenditure*, DGBAS, Executive Yuan, Taiwan (2016)	5.4349	0.6944
NX / Y	Ratio of total net exports to GDP	*AREMOS Taiwan Economical Statistical Databank*, Taiwan Economic Data Center	0.0685	0.0434
NX_O / Y	Ratio of net exports to OECD countries to GDP	*AREMOS Taiwan Economical Statistical Databank*	-0.0065	0.0670
NX_{NO} / Y	Ratio of net exports to non-OECD countries to GDP	*AREMOS Taiwan Economical Statistical Databank*	0.0750	0.0415
TFP	Total factor productivity index in industrial and service sectors, 2010=100	*Trends in Multifactor Productivity* (various years)	89.9973	8.8843
N^H / N^L	Ratio of the number of employed persons per household of the highest 20% to that of the lowest 20% households	*Report on the Survey of Family Income and Expenditure*, DGBAS, Executive Yuan, Taiwan (various years)	3.2081	0.9559
EMP_{SER} / EMP_{IND}	Ratio of the number of employees in the service sector to the number of employees in the industrial sector	*Yearbook of Manpower Survey Statistics*, DGBAS, Executive Yuan, Taiwan (various years)	1.3491	0.2767
LS_C / LS_M	Ratio of the number of employees with a bachelor's degree or higher to those with mandatory education or lower	*AREMOS Taiwan Economical Statistical Databank*	0.8918	0.7655
$CREDIT / Y$	Ratio of credit (loans and discounts) to the private sector by monetary institutions to GDP	*Financial Statistics Monthly, Taiwan District, ROC*, The Central Bank of China (Taiwan) (various years)	1.1117	0.3164
$SWEXP / GB$	Ratio of social welfare and social insurance expenditures to government budget. Because of lack of data for 1980, data on 1981 is used as a proxy.	*National Statistics*, Republic of China (Taiwan)	0.1239	0.0518

JUDITH LIU

Syracuse University	(315) 744-8939
Center for Policy Research	jliu61@syr.edu
426 Eggers Hall	https://jliu61.expressions.syr.edu
Syracuse, NY 13244	

EDUCATION

M.A. in Economics, Washington University in St. Louis, May 2014

M.A. in Economics, National Taiwan University, June 2012

B.A. in Economics, National Chengchi University, June 2010

FIELDS OF INTEREST

Public Economics, Labor Economics, Health Economics

RESEARCH

Publications

"The Long-Term Impact of Work-Hour Regulations on Physician Labor Supply," *Research in Labor Economics*, accepted

"The Differential Incidence and Severity of Food Insecurity by Racial, Ethnic, and Immigrant Groups Over the Great Recession in the United States," with A. Flores-Lagunes, H. Jales, and N. Wilson, *American Economic Review: Papers & Proceedings*, 2018, 108 (3): 379-383

"Economic Globalization and Household Income Inequality: An Empirical Study of Taiwan," with M.-Y. Lai and Z.-S. Liu (in Chinese), *Taiwan Economic Forecast and Policy*, 2019, 49 (2): 1-46

Research Papers

"Differential Exposure to Food Insecurity and Government Policy," with A. Flores-Lagunes, H. Jales, and N. Wilson

"The Effects of Trade Liberalization and Domestic Reforms on Income Inequality in a Middle-Income Economy: Evidence from Taiwan," with M.-Y. Lai and Z.-S. Liu

RESEARCH EXPERIENCE

Research Associate, Center for Policy Research, Syracuse University	Fall 2018–Present
Research Associate, Aging Studies Institute, Syracuse University	Fall 2016–Spring 2018
Research Assistant, Center for Policy Research, Syracuse University	Summer 2016
Research Assistant, Chung-Hua Institution for Economic Research	Fall 2011–Spring 2012

CONFERENCE PRESENTATIONS

Midwest Economic Association (MEA) Annual Meeting, St. Louis, MO	2019
Allied Social Sciences Associations (ASSA) Annual Meeting, Philadelphia, PA	2018
Agricultural and Applied Economics Association (AAEA) Annual Meeting, Chicago, IL	2017

FELLOWSHIPS, HONORS, AND AWARDS

Outstanding Teaching Assistant Award, Syracuse University	2017
Travel Grant, Syracuse University	2017–2019
Graduate Assistantship, Syracuse University	2014–2019
Maxwell Dean Summer Fellowship, Syracuse University	2014–2018
McDonnell International Scholarship, Washington University in St. Louis	2012–2014
Taiwan Ministry of Education Fellowship	2012–2014
Chung Hwa Rotary Educational Foundation Scholarship	2011–2012
Graduate Assistantship, National Taiwan University	2011–2012
Member of Phi Tau Phi Scholastic Honor Society	2010
Outstanding Graduate, National Chengchi University	2010
Academic Achievement Awards, National Chengchi University	2007–2010

TEACHING EXPERIENCE

Primary Instructor

Economic Ideas and Issues, Syracuse University	Summer 2017

Teaching Assistant

Public Economics, Syracuse University	Spring 2019
Program Evaluation (PhD-level), Syracuse University	Spring 2018
Intermediate Microeconomics, Syracuse University	Fall 2017-18, Spring 2019
Economic Ideas and Issues, Syracuse University	Fall 2014–Spring 2017
Innovation and Industrial Economics, National Taiwan University	Spring 2012
Statistics and Econometrics, National Taiwan University	Fall 2011
Principles of Economics, National Taiwan University	Fall 2011–Spring 2012

Academic Support Specialist in Economics

Office of Disability Services, Syracuse University	Fall 2017–Spring 2018

Teaching Training

Certificate in University Teaching, Syracuse University	Spring 2018

UNIVERSITY SERVICE

Selection Committee, Excellence in Graduate Education Faculty Recognition Awards, Syracuse University	2017, 2018, 2019
Presenter and Panelist, International Graduate Student Orientation, Graduate School, Syracuse University	2016, 2017, 2018
Treasurer, Taiwanese Graduate Student Association, Washington University in St. Louis	2013–2014
Presenter, Washington University Information Session	2013, 2014

REFERENCES

Alfonso Flores-Lagunes
(primary advisor)
Professor of Economics
Syracuse University
(315) 443-9045
afloresl@maxwell.syr.edu

Gary V. Engelhardt
Professor of Economics
Syracuse University
(315) 443-2703
gvengelh@syr.edu

Hugo Jales
Assistant Professor of Economics
Syracuse University
(315) 443-9230
hbjales@maxwell.syr.edu

Jerry Evensky
(teaching reference)
Professor of Economics
Syracuse University
(315) 443-5863
jevensky@maxwell.syr.edu

Elizabeth Ashby
(teaching reference)
Assistant Professor of Economics
Syracuse University
(315) 443-1354
eakerste@maxwell.syr.edu