THREE ESSAYS ON LABOR AND DEVELOPMENT ECONOMICS

By

YING-MIN KUO

Bachelor of Arts in Economics
Soochow University
Taipei, Taiwan
2008

Master of Arts in Economics
Oklahoma State University
Taipei, Taiwan
2010

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2020

THREE ESSAYS ON LABOR AND DEVELOPMENT ECONOMICS

Dissertation Approved:

Mehtabul Azam

Dissertation Adviser

Harounan Kazianga

Bidisha Lahiri

Brian Whitacre

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor Dr. Mehtabul Azam. He has been a wonderful mentor, research collaborator, and friend. I will always be thankful for his guidance when I felt lost in my PhD journey. I thank him for his support and advice on both research and on my career. It has been an honor to be his Ph.D. student.

I have appreciated my committee members, Dr. Harounan Kazianga, Dr. Bidisha Lahiri, and Dr. Brian Whitacre, for devoting their time to read my work and for their insightful suggestions. Their support enable me to complete my dissertation defense online during the coronavirus pandemic. They all really mean a lot to me.

My Ph.D. work would not be possible without the funding sources. I would like to thank the Fulbright Foundation for providing me with scholarship to gain academic and cultural experience in the U.S.A. I am grateful for Fellowship and Graduate Assistantship provided by the Department of Economics at Oklahoma State University.

None of this research would be done without my family and friends. A special thanks to my parents. Their encouragement and love gave me the self-esteem to overcome obstacles. I am also pleased to say thank you to my high school friends, Beryl, Erin, and Monica, who are always there when I need them.

In the end, I have cherished the friends I made at the OSU. I am grateful for time spent with them. I give special thanks to Kris and Sunny who accompanied me on my journey. This list of acknowledgments can only capture a small fraction of the people who supported my work. I send my deep thanks to all.

Name: YING-MIN KUO

Date of Degree: MAY 2020

Title of Study: THREE ESSAYS ON LABOR AND DEVELOPMENT ECONOMICS

Major Field: ECONOMICS

Abstract:

My dissertation comprises three chapters. The first chapter analyzes how economic shocks from the housing market impact the labor market and family formation decisions in the US. The second chapter explores the determinants of household fuel choices in India accounting for women bargaining and village infrastructure. The third chapter assesses impact of China's Belt and Road Initiative on the bilateral trades.

The first chapter assesses the impact of housing cycles in the USA on the family formation decisions. Pooling together information from datasets such as the Current Population Survey (CPS), Vital Statistics Natality data, the FHFA Housing Price Index (HPI), I construct outcomes at the Metropolitan Statistical Areas (MSA) level for various years, and exploit the variation in change in housing demand across MSAs to study the impact of housing demand shocks experienced by MSAs on the family formation outcomes in the MSAs. Recognizing the endogeneity of housing shocks, I use the magnitude of structural breaks in housing prices as instrument for housing shocks experienced by an MSA. The result support the hypothesis that increased employment opportunities during housing boom make men more marriageable but the impact of increased employment opportunities for women remain ambiguous because of countervailing positive and negative income and independence effect.

The second dissertation chapter investigates the fuel choice behavior of households using a large nationally representative panel data in India. Recognizing that the household choices are not independent over time, I use multinomial random and fixed effects models. The improvement over existing literature come from not only use of panel data, but also use of various women bargaining information and village level infrastructure. Women's education attainment and financial independence increase the plausibility of household using clean fuel. The existence of paved road is also an important determinant of household's decision in adopting clean fuel.

In the third chapter, I examine the impact of being connected to Silk Road railways on the 13 destination countries bilateral trade with China. To address the issue of endogenous placement of railways, I utilize the generalized synthetic control method. I find that the railway connections do not have significant impacts on either gross import or gross export of destination countries. The results are robust to alternative methods, across product categories, and different grouping of destination countries based on whether they share boundary with China.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

HOUSING BOOMS AND BUSTS AND FAMILY FORMATION OF YOUNG GENERATIONS

1. Introduction

Household formation has become a critical issue in the US due to significant changes in the household structure, such as an increase in the number of single-headed households (Black, McKinnish,and Sanders, 2003) and non-marital birth (Kearney and Wilson, 2018). Moreover, millennials tend to live with their parents or relatives and delay their transition to adulthood: leaving home, completing education, securing a job, marrying or cohabiting, and having children (Settersten and Ray, 2010; Sironi and Furstenberg, 2012; Sironi, 2018).[1] Local economic conditions should play a pivotal role in young people's decision making. For example, Schaller (2016) finds demand-induced changes in labor market conditions at the state level influence the fertility. Kearney and Wilson (2018) find PUMA-level fracking boom increased wages and jobs for non-college educated men, both marital and nonmarital birth rates increase, but marriage rates do not. Autor, Dorn, and Hanson (2019) exploit the differential impact of rising international manufacturing competition from China during 1990–2014 across commuting zones, and find that larger competition induced a negative local labor demand shock reducing employment and earnings of young adult males and co-

[1] Both young men and women since 1970s has postponed the transition and spent longer time to achieve each of them (Sironi and Furstenberg, 2012).

ontributed to reduction in marriage and fertility. In this paper, I investigate a hypothesis: housing booms and busts change young people's labor market and economic conditions, and the economic security affects their marriage and fertility decisions.

Following Charles, Hurst, and Notowidigdo (2018), I generate MSA-level housing demand shocks during 2000-2006, which is the sum of changes in local housing prices and changes in housing supply. Using the cross-MSA level variation in the housing booms, I study the impact of housing booms on MSA level labor market opportunities for all young men/women.[2] I then use a reduced-form model to analyze relationship between housing booms and demographic outcomes: marriage rates, fertility rates, and nonmarital birth rates. Admittedly, measurement error may exist in the housing demand shock and some latent factors that are associated with housing demand shocks may be omitted. To resolve plausible problems of endogeneity, I exploit cross-MSA variation in the magnitude of structure breaks of local housing price as an instrument variable (IV). As noted by others (Charles, Hurst, and Notowidigdo 2018; Shiller 2008; Mayer 2011; Sinai 2012) the variation in housing prices during the boom and bust derived from a speculative "bubble" and not from changes in standard determinants of housing values such as income, population, or construction costs (Shiller 2008; Mayer 2011; Sinai 2012), and hence the magnitude of structural breaks should be exogenous to unobserved local variables. The IV should not have a systematic relationship with pre-period observed variables. In the section 2.1, I show that the previous marriage rates and changes in marriage are not related to the IV.[3]

The housing cycles since 1990s in US has attracted extensive attention.[4] Strikingly, there is only a limited literature examining housing shocks on the family formation decisions, and most existing literature pay more attention to the effect of changes in housing

[2]Charles, Hurst, and Notowidigdo (2018) established that 2000-2006 housing booms increased employment opportunities and wage for young individuals without college degree. I extend their analysis to all young.

[3]For the fertility outcomes, although there is no relation between previous nonmarital birth rate and structural breaks, I fail to reject that IV is uncorrelated with previous fertility rate.

[4]In addition to a wider debate about the relationship between housing cycles and business cycles on the basis of macroeconomic analysis (e.g., Leamer, 2007; Ghent and Owyang, 2010), some papers focus on the micro-economic consequence of the real estate market fluctuation: saving and consumption (e.g., Skinner, 1994; Sheiner,1995; Campbell and Cocco 2007; Corradin and Popov 2015), household mobility (e.g., Engelhardt,2003; Ferreira, Gyourko, and Tracy, 2010), and labor supply (e.g., Farnham and Sevak, 2015; Rogers and Winkler, 2014; Laeven and Popov, 2016; Mian and Sufi, 2014; Charles, Hurst, and Notowidigdo, 2018).

costs—rents, housing prices, or mortgage. Rogers and Winkler (2014) examine whether housing market conditions, measured by the MSA-level housing prices, rents as well as foreclosure rates, during 2006-2011 affect the decision to live independently among people age 22-34. Through logit models, they find that housing prices has no significant impacts. OLS estimation results in Bowmaker and Emerson (2015) show a negative relationship between the burden of housing costs and marriage rate in US counties during 1970 to 1999. Using MSA-level CPS data, Farnham, Schmidt, and Sevak (2011) find fall in housing prices decreases homeowners' divorce risk relative to renters because high transaction costs make them locked into house/marriage.

With regard to fertility, more articles discuss about the connection between household income and fertility. Based on the model of fertility (Becker 1960; Becker and Lewis, 1973), an increase in husband's wage increase the demand of children through an income effect with the assumption that children are normal goods. Differently, an increase in woman's wage generates not only income effect but also substitution effect because of a higher cost of children. Also, considering a quantity-quality trade-off, household with higher income may pursue a higher quality of children instead of demand of children. That is, the negative relationship between income and fertility may exist. The puzzle with models is not completely answered by empirical evidence.[5]

Focusing on the wealth changes from the US housing market, Lovenheim and Mumford (2013) use probability models and the PSID data to analyze whether fertility decision is affected by changes in housing wealth at state or MSA levels measured by individual housing prices. When housing prices increase, the probability to have a child for homeowners age 25-44 increase by around 17%, which reflects housing wealth effect on homeowners' fertility decision. They also find no impact among renters. Dettling and Kearney (2014) analyze the effect of MSA-level housing prices during 1997-2007 on local fertility rates of women age 20-44. Compared to Lovenheim and Mumford (2013) studying the effect of housing prices on homeowners and renters separately, Dettling and

[5] Some empirical analysis observe a negative correlation between income and fertility, such as Jones, Schoonbroodt, and Tertilt (2010) and Lindo (2010), while some other find a positive relationship, such as Lovenheim and Mumford (2013).

Kearney (2014) adopt OLS and IV models and add an interaction term between housing price and ownership rate in the regression for distinguishing housing wealth effect and substitution effect. Due to the probability of reverse causality or unobservable variables that may affect both fertility and housing prices, they utilize the MSA supply elasticity as an instrument. They find a positive wealth effect as Lovenheim and Mumford (2013), but they also find a negative substitution effect: a rise in housing price decreases fertility rate because for potential homeowners higher costs of housing may decrease the demand of children with the assumption that children are normal goods.[6]

Different from these literature, I exploit changes in both prices and quantities of housing instead of only housing prices to capture changes in labor market conditions due to housing cycle and thereby the decision on family formation. Moreover, I consider men and women separately in the analysis of marriage rate as changes in economic conditions generate different impacts on young men and women. Improvement in employment and wages positively affect men's marriage (e.g., Gary, 1981; Gibson-Davis, 2009; Gibson-Davis, Edin, and McLanahan 2005, Sironi and Furstenberg, 2012) while the effect for women is ambiguous (Becker, 1965; Mincer, 1962). Women who have ability to be economically independent tend to delay marriage, and this is so-called women's independence effect (Oppenheimer, 1997). Nevertheless, a woman's income can lead a family to have a higher standard of living. The income effect makes women more likely to marry. Hence the effect of changes in economic condition on women depends on which effect dominates.

Using the data from the Current Population Survey (CPS), four panels of Figure 1 plot share of persons who reported being married separately for men and women in the 18-25 and 26-33 age groups. Marriage rate steadily declines since 1980. However, from around 1998, an upward deviation in the marriage rates from the trendline is observed among four groups. This is also the time when the housing boom started. At the end

[6] My fertility analysis is much similar to Dettling and Kearney (2014), but we have several differences. First, my analysis focus on a more recent housing cycle, the 2000-2006 boom and the 2006-212 bust, while they have a boom period between 1997-2006 and two bust periods, 1990-1996 and 2007-2009. Second, I study fertility rate of younger women aged 18-33 and restrict the sample by the condition that women's residence is occurrence place to remove the possibility of immigration. They analyze women age 20-44 instead. Moreover, I have a sample of 235 MSAs while they have 154 MSAs afetr merging all main dataset.

of boom and during the bust period 2007-2012, while the marriage rates for women age 26-33 gradually converge towards the trend, the marriage rates for the other three groups of people do not show convergence towards trend. Figure 2 plots the share of married for different birth cohorts in MSAs that belong to top tercile of the price increase between 2000-2006 and MSAs that belong to bottom two terciles of the housing price increase between 2000-2006.[7] As evident from the Figure, marriage rate is higher in MSA that experienced lower increment in housing prices during 2000-2006. It is worth noting that although marriage rates were declining among both group of MSA beginning 1970, some divergence in marriage rates between the two group of MSAs is evident around and after the 1976 birth cohort. Individuals born over 1975 and 1980 are my focus group because they were experiencing the 2000 housing boom when they were aged over 18. Figure 1 and Figure 2 provide initial assessment that the housing booms could have affected marriage rates and the effects may differ by gender.

Utilizing data from the National Center for Health Statistic (NCHS) and the vital statistics jurisdictions, Figure 3 shows the fertility rate for MSA grouped on the basis of terciles of housing price growth between 2000 and 2006 where the sample is restricted to mothers age 25-48 belonging to 1965-1987 birth cohorts. As evident from the Figure, there was a relatively larger increase in fertility rate for 1975-1980 cohorts living in the top-tercile MSAs. Moreover, how changes in economic condition affect nonmarital birth rate is also one of the interests in this paper. Rising rate of nonmarital birth has triggered great attention recently. Kalil and Ryan (2010) defines a household formed due to a nonmarital birth as a fragile family, which is relatively vulnerable to poverty and material hardship. Understanding the factors that may change the rate of nonmarital birth rate may help policy makers to know how to support this type of families. Figure 4 presents the difference of normarital birth rate between 1975-1980 cohorts living in the top-tercil MSAs and in the bottom-two-tercile MSAs is smaller. The rate of nonmatial birth over 1975-1980 cohorts in the top-tercile grows relatively slower. The housing boom

[7]Two series in the figure are marriage rates of people age 25-54 living in their birth state and belonging to 1965-1987 birth cohorts, which are constructed by the 1990 and 2000 censuses and 2005-2013 American Community Survey (ACS).

may cause a lower share of unmarried mother.

The findings of the paper are following. First, housing booms result into higher employment rate and a higher wage rate for young individuals age 18-25, while individuals age 26-33 only experienced a marginal increase in employment rate. Moreover, the impact of housing booms differs for men and women. While the housing boom increased employment for 26-33 aged men, it has no impact on 26-33 aged women. Though housing boom has similar impact on the employment opportunities for men and women in age 18-25, it resulted in much higher increase in average wages for men age 18-25. Second, I find that the housing booms increased the marriage among young men age 18-25 but is does not have significant impact on women's marriage rate in same age group. Moreover, I do not find any significant impact of housing boom on marriage rate of individuals age 26-33. Third, I find that housing booms led to increased fertility among 18-25 and 26-33 age group women, and decreased non-marital birth in both age groups. Fourth, during the 2006-2012 bust, housing booms continuously and positively impact marriage for men age 18-25, so the marriage rate at the end of the cycle does not return to the beginning of the housing cycle. I find new cohorts of women in the 26-33 age group response to the housing bust in a symmetric way: a lower marriage rate, a lower fertility rate, and a higher nonmarital birth rate.

My finding—housing booms generated positive economic shock for young men/women in age 18-25,but they only affected the marriage rate of men positively—supports the idea that improved labor conditions for men make them more marriageable. Moreover, as noted in the literature, the improved labor market conditions for women would lead to two counteractive effects: income effects and independence effect. As a result, I do not see impact on marriage rates for women in age 18-25. Increased fertility stemming from positive labor demand shock is also not surprising. My results are in concordance with Autor, Dorn, and Hanson (2019) finding that increased import competition from China that led to negative demand shock for blue color men workers decreased marriage rate and fertility rate.

This paper contributes to the literature in following ways. First, to the best of my knowledge, this paper is the first paper that examine this channel: the impact of housing boom/busts on family formation decisions though changes in labor market conditions. I extend Charles, Hurst, and Notowidigdo (2018) work that focus on the impact of housing booms/busts on the labor market outcomes for individuals in age 18-25 without any college education to all men and all women ages 18-25 and 26-33. Second, I focus on the outcomes of much younger generations, age 18-25 and 26-33, who are sensitive to economic fluctuation (Sironi, 2018). Moreover, I analyze the effect of a more recent housing cycle. Selecting the 2000-2013 cycle because recent literature have consensus that changes in housing demands are the main source of the 2000s housing booms (e.g. Shiller, 2008). This enables me to utilize changes in housing demand to measure the housing boom and establish the channel.

The remainder of this paper is organized as follows. The second section describes the data, construction of housing demand shock, construction of the instrument variable, and empirical strategy. Section 3 presents the results for marriage rate and fertility rates, and section 4 concludes.

2. Data and Empirical Approach

2.1. Local Housing Demand Shocks

Following Charles, Hurst, and Notowidigdo (2018), I construct the housing demand shock across MSAs during the 2000-2006 boom for analyzing its effects on household formation. The construction of the shock is based on recent consensus that the housing boom over 2000s mainly came from the changes in housing demand (e.g., Shiller 2008). Moreover, the difference in local housing price appreciations among MSAs during the boom is caused by both different changes in local housing demand (e.g., Ferreira and Gyourko, 2011) and in local housing supply elasticity (Mian and Sufi 2011). Using a log-linear model of housing demand and housing supply, a local housing demand shock, ΔH_m^D, is in the following form:

$$\Delta H_m^D = \eta_m^D \Delta P_m + \Delta Q_m \tag{1}$$

where ΔP_m is the change in the log of local housing prices in MSA m, η_m^D is the price elasticity of housing demand, and ΔQ_m is the change in log of quantity of new housing.. Similar to Charles, Hurst, and Notowidigdo (2018), I take the existing estimates of the elasticity of housing demand in the literature that suggest $\eta_m^D \approx 1$ to create proxy for demand shock.[8] The proxy of housing demand shock is equal to the sum of log difference in local housing prices and log difference in quantity of housing in the MSA m. Both housing prices and housing supply channels should influence local labor market in the theory (Mian and Sufi, 2014; Charles, Hurst, and Notowidigdo, 2018). A rise in housing prices can increase household wealth and liquidity, so it can stimulate local consumption and employment. An increase in housing supply reflects a higher labor demand in the local construction industry.

Using the MSA-level data of local housing price provided from Federal Housing Finance Agency (FHFA), I compute annual local housing prices by averaging quarterly series of local housing price index (HPI) and manually match FHFA metro areas to the ACS metro areas. I adjust nominal HPI to real HPI using CPI-U published by BLS. The base year of HPI is 1995. The way FHFA metro area are matched to ACS metro area is shown in Data Appendix. I use numbers of housing permits as a proxy for the housing supply, i.e. the quantity of new housing.[9] 275 MSAs can be matched after I merge ACS data with the FHFA and permits survey data.[10]

As noted by Charles, Hurst, and Notowidigdo (2018) and evident from Appendix Figure A1, changes in housing prices and in housing permits during 2000 to 2006 are positively correlated, which suggests role of housing permits in the local housing demand shocks. Figure A2 plots the trend in the demand measure at 10th, 50th, and 90th

[8] Charles, Hurst, and Notowidigdo (2018) took the average of two widely-cited estimates of the housing demand elasticity in the literature: 0.7 from Polinsky and Ellowood (1979) and 1.2 from Houthakker and Taylor (1970).

[9] I thank Charles, Hurst, and Notowidigdo (2018) for sharing housing permit data. Detail of permit data is given in Charles, Hurst, and Notowidigdo (2018).

[10] Charles, Hurst, and Notowidigdo (2018) also has 275 MSAs.

percentiles during 1990-2012. There is a significant changes at 90th percentile over boom and bust periods comparing to changes at median and 10th percentile. Similar to Charles, Hurst, and Notowidigdo (2018), I exploit the large variation in housing demand across MSAs. The standard deviation across MSAs in the 2000-2006 is 0.464.[11]

Although changes in the housing demand shock is exogenous, potential measurement error in construction of the shock leads to attenuation. In addition, biases can also be caused by omitted unobserved factors that are correlated with the housing demand shocks, such as other labor market factors or marriage policies. To address potential endogeneity issues, I adopt a Two Stage Least Squares (2SLS) estimation strategy using an Instrument variable (IV). Following Charles, Hurst, and Notowidigdo (2018) and Ferreira and Gyourko (2011), I use the magnitude of sharp changes in housing prices in each MSA over 2000-2006 as an instrument. The magnitude of sharp change in housing prices captures exogenous variation in housing demand shock resulted from speculative bubbles instead of the traditional determinants of housing prices (Ferreira and Gyourko, 2011; Charles, Hurst, and Notowidigdo, 2018). The fundamental concept behind is built on the consensus that dramatic housing boom since around 2000 resulted from a high expectations of future home values. Higher expected housing price level led to more speculative housing investment (Shiller, 2008).[12]

To find the sharp changes, I use quarterly local housing prices between 2000q1 and 2005q1 and find a local structure break maximizing the R^2 of the following MSA-specific OLS regression:

$$P_m^H(t) = \omega_m + \tau_m\, t + \lambda_m(t - t_m^*)1\{t > t_m^*\} + \xi_{m,t} \tag{2}$$

where $P_m^H(t)$ is the log of local real HPI in MSA m in year-quarter t. t_m^* is the date of

[11] Standard deviation is 0.55 in Charles, Hurst, and Notowidigdo (2018). A lower standard deviation in this paper may mainly be caused by the difference in base years of HPI. The base year of HPI is 1995 in my analysis while 1980 in their work. I also rebase the CPI which used to make raw HPI a real variable.

[12] In Appendix Figure A3, I plot the quarterly housing prices for six sampled MSAs between 2000:I and 2005:IV. For the three MSAs in the left panel, the evolution of prices over time has been smooth, while for the three MSAs in right panel price series changed "sharply" at some point in the 2000s, suggesting the influence of some factor different from smooth changes in fundamentals, such as the effect of a speculative bubble (Charles, Hurst, and Notowidigdo, 2018). Importantly, the sharp change for different MSAs differ across MSAs, and I use the variation in magnitude of these sharp changes as Instrument.

the MSA-specific structure break between 2001q1 and 2005q1.[13] λ_m is the magnitude of the structure break in MSA m. τ_m captures a MSA-specific time trend before the break.

2.2. The Effect of Housing Boom on Family Formation

My interest is to examine the impact of housing shocks on family formation decisions and labor market outcomes of young men/women. For this, I estimate the following first-difference regressions:

$$\Delta S_{mt} = \beta_0 + \beta_{FD}\widehat{\Delta H^D_{mt}} + X_{mt}\beta_3 + u_{mt}. \tag{3}$$

where $\widehat{\Delta H^D_{mt}}$ is the change in housing demand as estimated by equation (1) and ΔS_{mt} is the change in outcome of interest over time t and $t+k$. The first differenced model account for any time invariant MSA-level factors. X_{mt} include MSA-level baseline controls that could cause differential trends in outcome variable: log of the age 18-55 population, the employed with a college degree, foreign born share, female employment rate in 2000. β_{FD} is the parameter of interest.

The data for labor outcomes and marital status are constructed from the Census/ACS samples. I restrict the data to people age 18-33 living in the birth state so as to exclude the potential effects of migration. I also exclude people living in the group quarters. In my analysis, I focus on outcomes measure in 2000, 2006, and 2012. While the labor/marriage outcome variables in 2000 is computed from 2000 census, I use pooled ACS data over 2005-2007 and over 2011-2013 respectively to compute the outcome variables for year 2006 and 2012 for a more precise analysis at MSA-level.[14]

The fertility data is from the Vital Statistic Natality Files. Data of population by gender and age are from the Surveillance, Epidemiology, and End Results (SEER) database. The sample is restricted to woman in age categories 18-25 and 26-33 whose residence is occurrence place. I map the data by the state and county Federal Information Processing

[13] Following Andrew (1993), I use 15% as of the left and right trimming percentage of sample period.

[14] In this paper, I mainly use METAREA code to merge data. The code is not available in 2012 and 2013 so I manually match MET2013 code in these two years to METAREA.

Standards (FIPS) codes to MSAs. I also exclude the MSAs with less than 5 natality data. There are 235 MSAs remained. The fertility rate in year t at MSA m is equal to the number of local birth over women population in a specific region. The nonmarital birth rate is calculated by the number of local nonmarital birth over regional total birth.

Table 1 reports summary statistics of variables used in the analysis. The main sample for housing market and marriage includes 275 MSAs. The first row of Table 1 shows MSA housing demand shock rises by around 50% on average over the 2000-2006 boom period. The housing demand shock contains price effect and supply effect. Second and third rows displays that in the 50-percent increase in housing demand, 34.7 percent comes from the change in housing prices during the housing boom, and the remaining part is from the increase in housing supply (16 of 50 percent change). Arranging MSAs according to price increase over 2000-2006, the housing prices increased by 71.9% for the 90th percentile MSA and by 5.9% for the 10th percentile MSA. The second section of Table 1 are the summary statistics of control variables. The remaining section has key outcome variables of labor market, marriage, fertility rate, and nonmarital birth rate.

3. Results

3.1. Changes in Labor Outcomes from Housing Booms

Charles, Hurst, and Notowidigdo (2018) finds that housing booms improve the employment rate and average wage rate for young people age 18-25 without college degree during booms, increasing the opportunity cost of acquiring a college degree. Extending the analysis to all education groups, I assess whether housing booms enhance the labor market conditions and benefit all young men and women in two age categories, 18-25 and 26-33. Table 2 shows the first-stage results that suggest a strong relation between the size of structure break and changes in housing demand shock. Table 3 presents both OLS and 2SLS results in Panel A and panel B respectively.[15]

[15] Table 3 only presents the effect of 2000-2006 housing demand change on labor market outcomes. The full result of OLS estimate is shown in Table A1.

OLS results suggest positive and statistically significant impacts of change in housing demand over 2000-2006 on both change in employment and change in average real wages for people age 18-25 over 2000 to 2006. In panel A, OLS results suggest that 100 log points increase in housing demand lead to 2.2 percentage points increase in employment for all young in age 18-25, and 2.5 and 3.3 percentage points increase in employment for men and women in age 18-25, respectively. Moreover, real wage rate increases by 6.4, 8.1, and 4.0 percent for all, men and women age 18-25 as a result of 100 log point increase in housing demand. Thus the OLS results suggest improved labor market opportunities for young in age 18-25, and importantly improvement is observed by both genders. There are smaller positive impacts on people in the age group of 26-33. A 100 log points increase in housing demand shocks induce 1.5 percentage point and 3.4 percent increase in employment rate and average wage rate for people in this 26-33 age group. While, OLS results suggest positive impact of change housing demand for men in age 26-33, I could not reject the null of no impact for women in the 26-33 age group.

In panel B of Table 3, the 2SLS point estimates are generally larger than OLS estimates indicating dominance or presence of attenuation bias caused by measurement error. The 2SLS estimates suggest that a 100 log point increase in housing demand increased employment rate of all young, young men, and young women in age group 18-25 by 3, 3.6, and 4.1 percentage points respectively. At the same time, the same increase in housing demand increased wages by 8, 10.7, 3.2 percent for all, men, and women in age group 18-25. A one standard deviation change in housing demand shock is 0.49 across MSAs. In other words, a one standard deviation (SD) increase in housing demand causes 1.4, 1.8, and 2 percentage point increase in employment for all people, men, and women in age group 18-25 respectively.[16] Similarly, a one SD increase in housing demand increase the wages of all, men, and women in age group 18-25 by 4, 5.2, and 1.5 percent, respectively. The IV results suggest a much weaker effects of housing demand increase

[16]The percentage point change is calculated by the coefficient from the table times the standard deviation of change in housing demand shock. For example, a one standard deviation increase in housing demand results in $3.6 \times 0.49 = 1.8$ percentage point increase in employment for men age 18-25.

for age 26-33. Changes in housing demand shocks by 100 log points make employment rate for people age 26-33 increase by 1.9 percentage points which is higher than the OLS estimate in panel A. Importantly, there is no statistically significant impact of wages of 26-33 age group. While the impact on employment of 26-33 aged men is statistically significant, the impact on employment of 26-33 aged women is statistically insignificant. There is no impact on the wages of either gender in 26-33 age group.

The main measure of housing demand combined both change in housing prices and change in housing supply, and assumes that the effects of each of the two components are similar. I examine whether the effect of housing demand mainly comes from one of them by considering ΔP_k and ΔQ_k as separate variables in OLS model. The results are reported in Appendix Table A2[17] As evident from the OLS results, both components of housing demand have positive and fairly similar effects. That is, increases in housing prices and housing supply increase employment and wages for young people. Moreover, to assess whether housing demand shocks impact the local labor market differently because of distinct level of local housing supply elasticity, I include the local housing supply elasticity estimates from Saiz (2010) and its interaction term with the housing demand shock in the regressions. In 2SLS models, I have two IVs: the size of structure breaks in housing prices and its interaction term with the supply elasticity. I do not find differential effects of housing demand changes due to different supply elasticities. These results support that the effect of changes in housing prices and in supplies are similar, so the combination of both changes is exploited in the rest of the analysis in this paper.

Overall, the results of labor market outcomes confirm that the housing boom boosts labor market conditions especially for younger age group, 18-25-year-olds. For people age 26-33, the improvement in labor market conditions is mainly in men's employment.

[17] I could not implement the IV strategy as I only have one instrument variable

3.2. Changes in Marriage Rate from Housing Booms

As discussed in the last section, main mechanism—enhancement in young people's economic conditions—through which housing booms effect family formation decisions exist. Having established that housing booms may generate an income/wealth effect for young individuals, I move to the reduced form effects of housing booms on marriage outcomes. In this section, I focus on marriage rate as well as the share of other marital status, including cohabitation, divorce, and never married. If the shocks indeed change their decision in starting a family, some shares of these status at the aggregate level may be affected.

Panel A of Table 4 shows the IV impacts of housing booms on various measures of marital status for all 18-25-year-olds.[18] The IV estimate is statistically significant for share of married outcome at 10% significance level, and there is no statistically significant effect on other outcomes such as share of divorced, share of cohabitation, or share of non-married. A one standard deviation increase in housing demand shock increases marriage rate by 0.8 percentage points. Panel B and Panel C of Table present the results for men and women, separately. The impact of housing boom on marriage rate for men is statistically significant. Although the sign of the IV estimate for share of married outcome is positive for young women also, I cannot reject the null of no effect.

Moreover, I find no effect on share of cohabitation, which includes same-sex and different-sex cohabitation. In terms of divorce rate, there is only modest positive impacts for men age 18-25 (0.1 percent). Noticeably, there is no significant effect of housing shock on marital outcomes 26-33-year-olds. Considering that home ownership may affect their marriage status based on the literature, such as Farnham, Schmidt, and Sevak (2011), the last column extends baseline model of marriage rate with an extra control, home ownership rate. The result is almost as same as the baseline model.

Pooling together marriage outcomes (Table 4) and labor outcomes results (Table 3),

[18] Table 4 only presents the effect of 2000-2006 housing demand change on marriage status. The full marriage result of OLS estimate is shown in Table A3.

although housing boom increased the labor opportunities for both young men and women in age 18-25 age group, the housing boom only increased the share of married for young men and not for young women. As discussed earlier, theoretical predictions in the literature suggest that while the increased labor opportunities make men more marriageable, the impact of increased labor opportunity for women in ambiguous because of opposing independence effect and income effects.[19]

3.3. Changes in Fertility Rate from Housing Booms

In this section, I look at whether the birth outcomes among people age 18-25 and 26-33 changed as a result of housing booms. The fertility results are presented in Table 5. The IV estimates suggest that fertility rate for women in 18-25 age group increases by 1.2 percentage points in response to a one standard deviation increase in housing demand shock, whereas fertility rate for women in age 26-33 group increases by a marginally higher 1.4 percent. In contrast, nonmarital birth rates decrease during housing booms for both age groups. A one standard deviation increase in housing demand shock reduces nonmarital birth rate by 2.6 and 1.4 percent for 18-25-year-olds and 26-33-year-olds respectively.

Connecting the fertility result with Table 3, younger people age 18-25 are more sensitive to the housing boom, but their changes in fertility rate are smaller during the boom comparing to older age group even though this is not a big difference. This can be explained by the literature that, similar to the analysis of marriage rate, a better labor market condition usually increases the possibility to have a baby for men, yet the effect for women sometime would offset by opportunity costs of working. Table 3 suggests housing boom has impact on men in the 26-33 age group but not women in the same age group, while it has influence on both men and women in the 18-25. Hence, the results in Table 5 implies the coefficients for people age 18-25 capture the offsetting effect, so it is

[19] I also explore the effect of housing demand shocks on marriage rates of homeowners and renters separately. The results are shown in the Appendix Table A4. A caveat here is that homeownership/renter rates in MSA might itself be affected by change in housing prices over time introducing a bias in the measurement of outcome variables. The housing demand shocks still have a significant impact on the marriage of younger homeowners age 18-25 but no impact on the marriage of homeowners in the older age group.

relatively smaller than the effect for people age 26-33.

The impact of economic shocks on fertility is ambiguous. The effect of improvement in economic conditions on fertility may reflect a net effect involving wealth effect and substitution effect. The wealth effect implies children are normal goods. When labor market conditions has improved, extra wealth brings a higher fertility rate. However, higher return in working means a higher opportunity costs of parental time, and leads a lower fertility rate. A few papers find local adverse economic shocks cause reduction in birth rate (Ananat, Gassman-Pines, and Gibson-Davis, 2013; Autor, Dorn, and Hanson, 2019) and an increase in unmarried mothers (Autor, Dorn, and Hanson, 2019), while Kearney and Wilson (2018) finds a positive economic shocks make both fertility rate and nonmarital birth increase. My fertility results that suggest positive demand shock increased fertility is in concordance with Autor, Dorn, and Hanson (2019) and Kearney and Wilson (2018) as positive demand shock boosts birth rate. However, my results on non-marital births are in contrast to Kearney and Wilson (2018). While I find that increased labor demand reduced non-marital birth, Kearney and Wilson (2018) find that local-area fracking boom that increased wages and jobs for non-college educated men increased both marital and nonmarital birth rates. I think of two possible reasons for differences. First, I look at all young individuals, while Kearney and Wilson (2018) focus on non-college educated. Second, the nature of demand shock in my case have a much wider impact compared to fracking boom which is more likely to have impacts on non-college educated groups.

Of course, the birth rate may not precisely capture the time young people decide to have a baby because the decision may be made before the birth year. There may also be a delayed effect on birth rates. Young parents should feel they are better off and then decide to have a baby. Therefore, I change the outcome period to 2000-2007. The estimation is shown in column (3) and (4) in the Table 5. The result is rather consistent with the baseline result. If I modify the period to 2001-2007 in column (5) and (6), the effect is smaller and less significant. Shocks during housing boom seem to affect only fertility rate for people 26-33, but the sign of coefficients remain the same. In addition, Dettling and Kearney (2014) analyzing the effect of housing costs on fertility rate find the home ownership has significant impact on fertility rate. If a household have a house, the housing booms should bring extra wealth to this family, and this may lead to a higher birth rate. The omission of home ownership may cause omitted-variable problems. Since, Vital Statistics data do not contain ownership information, I add home ownership rate in MSA in 2000 from the ACS data as an extra control variable in the model. In the column (7) and (8), the results are similar to results in the baseline model.

Another way to measure birth rate is to exploit the year of conception. I use the month of birth and the length of gestation to find the approximate year of conception, so I can calculate the share of conception in a year. The results for this new birth rate are presented in Table A5 in the Appendix. There is no significant impact for women in two age categories, but the signs of coefficients remain the same.

3.4. Placebo Analysis

In this section, I examine whether the results in the previous sections are indeed capturing the causal effect of housing booms, or the regressions may simply be capturing the pre-existing trend. Using the baseline model, Table 6 presents the results of placebo tests for marriage outcome. I basically examine whether the 2000-2006 change in housing demand predicts the growth in average marriage rate from earlier period 1990-1996. The marriage rates for the 1990 and 1996 are computed from the CPS data as annual ACS data are not available over 1990-2000. As evident from the Table, IV estimates are statistically insignificant for both age groups and genders considered. Although this does not prove that my main results are not driven by time invariant factors that line up with variation in the housing demand or magnitude of structural breaks, but increase confidence in the main results as I can rule out any pre-existing trends in the outcome.

In Table 7, I present the placebo tests using change in fertility rate and nonmarital birth rates over 1990-1996. Although there is no correlation between pre-boom change in fertility outcome and housing boom for women in age group 18-25, I find a negative correlation between fertility and housing booms for women in age group 26-33. It is important to note that the correlation between housing booms and 1990-1996 change in fertility rate is negative, however, in my main results I find positive impact of housing boom on fertility rate. This suggests that my main fertility results underestimate the impacts on fertility for women in 26-33 age group.

3.5. Family Formation and Housing Busts

In this section, I examine the effect of shock from housing busts in the 2006-2012 period as well as entire 2000-2012 period spanning the entire boom and bust cycle. I use the fact that MSAs that witnessed larger booms during 2000-2006 period also experienced larger busts during 2006-2012 (as evident in Figure A5). By correlating the housing boom during 2000-2006 to change in outcomes over 2006-2012, I check whether the size of housing boom has a persistent effect or the effects of housing booms were reversed for new cohorts during the housing busts.

Column (1) in Table 8 is the baseline result. The shock from housing booms has a persistently positive impact on 18-25-year-olds during housing bust. In other words, during economic downturn, younger generations may still have higher marriage rate, this is different from my expectation. However, it generate a reversed effect for people age 26-33, especially for women. The results of the entire housing cycle seem to align with the initial assessment of Figure 1. For all, men, and women in the 18-25 age group, marriage rates at the end of cycle positively deviate from the 2000 level when the housing boom started while there is a convergence in marriage rate for women age 26-33.

To understand the plausible reasons of the persistent effect on the 18-25-year-olds, I firstly check if this effect is disappeared over time, because they might have made decision to marry at the end of the housing boom. I investigate the outcome by two sub-periods of housing bust, 2006-2009 and 2009-2012, and by the period from 2008 when the Great Depression starts to 2012. The size of housing boom has no or very small impact on marriage rate of 18-25-year-olds. I also looked at the effect of housing demand shock with extra control—the share of ownership rate at the beginning of the housing bust, but the results are similar to the baseline model.[20]

Table 9 presents the results for the fertility rate during the bust period. The size of housing boom during 2000-2006 has no persistent impact on fertility rate and non-marital birth rate for women in 18-25 during housing busts. However, we can see a reversed effect for women age 26-33 during the housing busts. In response to an adverse economic condition during the housing bust, the fertility rate of women in the older age group is lower, but the ratio of nonmarital birth increases, which is consistent with Autor, Dorn, and Hanson (2019). In fact, Schaller (2016), who use state level data to check if children are normal good as assumed in the Becker's model, find that the fertility rate is procyclical. My result also supports the idea of procyclicality in fertility rate. Looking at the entire cycle, I found no significant difference between the 2000 birth/nonmarital birth rate and the 2012 birth/nonmarital birth rate.

4. Conclusion

This paper begins with preliminary assessments that the recent housing cycle may affect household formation decisions. Considering a consensus from literature that the 2000 housing boom is mainly resulted from changes in housing demand, I utilized the changes in housing demand shocks to measure the size of housing booms and adopt the latest instrument for the booms—the size of structural break in housing prices during the housing booms—to identify the causality of housing demand shocks on labor market conditions and household formation decision.

This paper provides empirical evidence that housing booms improve labor market conditions for young people, and this economic security from housing booms lead to an increase in both marriage rate and fertility rate and a lower nonmarital birth rate. The results indicate that people in the younger age group, age 18-25, are more sensitive to

[20] The result is not shown.

housing booms, and this gives us an explanation to a significant increase in marriage rate for only 18-25-year-olds. This paper also find that during housing busts, 26-33-year-olds have a lower lower marriage rate. The increase in fertility rate and the decrease in nonmarital birth rate as result of the housing boom reversed during the housing bust. The results are in concordance with the Becker's model and empirical results from Autor, Dorn, and Hanson (2019).

The findings in this paper add to the literature by showing the effects of housing market fluctuation on labor market and thereby marriage and fertility decisions. These can help policy makers to understand how young people response to housing cycles and how policy can be formulated to support the households when they are facing economic fluctuations especially downturns. Admittedly, why the housing boom continuously increase marriage rate for the 18-25-year-olds during the housing bust is not fully explained in this paper. Future research could continue to explore the possible reasons.

CHAPTER II

HOUSEHOLD COOKING FUEL CHOICE IN INDIA, 2004-2012: A PANEL MULTINOMIAL ANALYSIS

1. Introduction

Around 3 billion people cook using polluting open fires or simple stoves fuelled by kerosene, biomass (wood, animal dung and crop waste) and coal (WHO, 2018).[21] Using traditional fuels, such as firewood, charcoal, or agricultural waste, produces both indoor and outdoor air pollution and threatens public health (Bruce et al., 2000; Smith, 2000; Alem et al., 2016). Collecting traditional fuels is time-consuming, and women and children are often responsible for this job. Thus use of traditional fuels reduces their time for studying or doing other productive activities (Burke and Dundas, 2015). For economic growth, it is crucial to replace traditional fuels with modern fuels, like electricity, kerosene, or liquefied petroleum gas (LPG) (Kaygusuz, 2011). Not surprisingly, the issue of household transition from traditional fuel to clean fuel has received considerable attention from both researchers and policymakers.[22] A number of studies published over the past three decades investigate the factors driving the transition. Muller and Yan (2018) provide a recent survey of the literature. As argued by Muller and Yan (2018), even though some studies are merely based on simple descriptive statistics, one can see the emergence of econometric methods to quantify the patterns and factors of household fuel use. Majority of the existing liter-

[21] https://www.who.int/en/news-room/fact-sheets/detail/household-air-pollution-and-health

[22] Cooking fuels also have attracted increasing interest over the years because fuel wood harvesting has caused extensive deforestation.

ature on fuel-transition is based on cross-section data, and only recently, researchers have used panel data to have more credible estimation strategy (Alem et al., 2016; Choumert-Nkolo et al., 2019).

In the literature on fuel transition from traditional to clean fuel, "energy ladder model" is quite popular (Hosier and Dowd, 1987; Leach, 1992; Leach and Mearns, 2013; Van der Kroon et al., 2013). "Energy ladder model" states that the households would move along the energy ladder when they receive higher income or social status. However, households are usually unable to get rid of traditional fuels completely because of cost consideration, culture preference, or supply side considerations (Masera et al., 2000; Alem et al., 2016). It has been noted that multiple fuel use constitutes the rule rather than the exception in many urban and rural areas of developing countries, and the use of multiple fuels is described as "fuel stacking" (Heltberg, 2004; Masera et al., 2000; Ruiz-Mercado and Masera, 2015).

In this paper, I use a nationally representative panel data from India, India Human Development Survey (IHDS), to examine determinants of cooking fuel choice in urban and rural India, separately. I use multi-period multinomial logit model of fuel choice that account for unobserved time and choice invariant household heterogeneity. In addition to the usual household characteristics and prices of different fuels, the richness of the IHDS data allows me to control village level infrastructure, and women's bargaining power.

Given the population size and development stage of India, interest in households' transition from traditional to clean fuels is not new, and there exists significant literature on this topic. However, the existing literature on India use either a single or multiple cross-section data. Gangopadhyay et al. (2003) estimate a multinomial logit to model household main fuel (both cooking and lighting fuels together) choice separately for rural and urban Indian households using one round of Consumer Expenditure Survey collected by National Sample Survey (NSS) in 1999-00. Farsi et al. (2007) also use one round of NSS cross-section data collected in 1999-00 to examine household cooking fuel choice in urban India. Compared to Gangopadhyay et al. (2003), they use an ordered discrete choice framework and focus only on cooking fuels for urban households. In comparison to above mentioned two studies that use single round of cross-section data, Viswanathan and Kumar (2005) and Cheng and Urpelainen (2014) use multiple rounds of NSS cross-section data. Viswanathan and Kumar (2005) use three rounds of NSS data collected in 1983, 1993-94, and 1999-00 to descriptively document expenditure on different types of fuel and share of clean fuel in total fuel expenditure in rural and urban India and Indian states. Cheng and Urpelainen (2014) use two rounds of NSS data collected in 1987-88 and 2009-10, and discuss fuel-stacking behavior for lighting and cooking in India. They use a two-stage model: in the first stage they estimate a probit model for use of any modern fuel, while in the second stage they examine whether households engage in fuel stacking behavior conditional on using any modern fuel. They find the stacking of LPG and traditional biomass has grown rapidly in India over 1987 and 2010. They also find that although the household income has a robust negative effect on cooking fuel stacking in 1987, it has a positive effect in 2010. They speculate that with the dramatic rise in the use of LPG in 2010, fuel stacking has become so common that even relatively wealthy households now engage in it.

It is worth noting that the dynamic fuel stacking behavior cannot be observed in cross-section data (Alem et al., 2016). Moreover, the covariates available in the NSS data, that most of Indian literature on fuel choices is based on, are limited. For example, the NSS data do not contain any information that can be used to measure the bargaining power of the women in the household except education of individuals. Similarly information on village infrastructure is not available in the NSS data.[23]

This paper contributes to existing literature in the following ways. First, unlike the

[23] Moreover, the NSS data report expenditure on consumption of energy (fuel, light and household appliances) during the last 30 days. Thus it does not distinguish between cooking and lighting. Cheng and Urpelainen (2014) reduce the dimensionality of the lighting fuel choice to that of kerosene and electricity, while focusing on liquefied petroleum gas (LPG) and biomass on the cooking part. While for most fuels the primary use for cooking or lighting is distinct, fuels such as kerosene could be used for both. According to 68th round of NSS data collected in 2011-12, about 2.7 percent of households reported kerosene as the main source of cooking while only 0.17 percent of households report electricity as main source of cooking. The 2011 Census data suggest that 2.9 percent of the households use kerosene as main source of cooking, while only 0.1 percent of households use electricity as main source. Nonetheless, even if kerosene might not be the main source of cooking, it could be a supplementary source of cooking. For example, 2011 IHDS that asks whether households use kerosene, and for what purpose, about 27 percent households responded using kerosene for either cooking or cooking and lighting both.

existing literature on India (Cheng and Urpelainen, 2014; Farsi et al., 2007; Viswanathan and Kumar, 2005; and Gangopadhyay et al., 2003) that is based on cross-section NSS data, I use a panel data and take account of household heterogeneity using a multinomial random effects model. [24] The advantage of panel data is that it relaxes the assumption — multiple observations within a choice are independent (Alem et al., 2016). Second, distinct from existing literature on India, I also look at the bargaining power of women using multiple indicators available in my data. Third, for rural households, I examine the impact of village level infrastructure on fuel choice.

The main findings of the paper are following. First, there exist substantial differences across urban and rural areas. The household characteristics that are positively associated with the use of clean fuel (e.g. income, education) also increases the use of fuel stacking in rural areas but have no significant effects in urban areas. This suggests that a clean-break with the use of traditional fuels is unlikely to be in occurring rural areas, while more probable in urban areas. Rural households are more likely to go through stages where they shift to mixed fuel and later on to clean fuel. Second, I find that access to paved road is an important determinant for rural households adopting clean fuel and distance to the nearest town is not important. Third, I also find evidence of social spillover effects in rural areas: households residing in villages that reported clean (dirty) fuel as their main source are more likely to use clean (dirty) fuel and less likely to use dirty (clean) fuel. Fourth, the bargaining power of women that is associated with economic status (e.g. education or economic freedom) leads to increase in probability of household using clean fuel. Fifth, I find considerable impact of LPG prices on the probability of use of clean fuel in urban areas, but no significant impact for rural households.

The rest of the paper is organized as follows. Section 2 details the data, Section 3 provides the empirical strategy, Section 4 presents the results, and Section 5 concludes.

[24] Recently, Mekonnen and Köhlin (2009) and Alem, et al. (2016) use multinomial logit model with random effects to study fuel choices in Ethiopia, Choumert-Nkolo et al. (2019) use multinomial logit model with random effects to study fuel choice in Tanzania.

2. Overview of Data

2.1. Data Description

I use two waves of large scale India Human Development Survey (IHDS) collected in 2004-05, and 2011-12 (henceforth, 2004 and 2011, respectively). The IHDS are nationally representative and were collected jointly by National Council of Applied Economic Research at Delhi and the University of Maryland (Desai et al. 2005; and Desai and Vanneman, 2015).[25] The 2011 IHDS collected information on 42,153 households (27,580 rural and 14,573 urban). Out of these 42,153 households, 40,018 households were also surveyed in the 2005 IHDS. I use only those households that were surveyed in both rounds. I further drop 653 households who do not report using any cooking fuel. Thus my final data contains a balanced panel of 39,365 households (26,927 rural and 12,438 urban).

The IHDS data contain several socioeconomic information at the household and individual level. The data have information on both household consumption and income, and information on each type of fuel used in the households. Unlike NSS survey that asks for main cooking fuel, IHDS contain detailed energy module where respondents were asked detailed questions about their use of all energy sources. A separate village and women module were also implemented as part of IHDS. The village module contains information of village infrastructure and prices of different fuels. I use prices of different fuels at the village level collected in village module. Since the prices were not available for urban areas, I import district level fuel prices in urban areas from NSS 61st and 68th rounds of consumption expenditure. The NSS 61st and 68th rounds were collected in 2004-05 and 2011-12, and overlap with sample period of IHDS. NSS data do not collect prices at town/village or district levels, hence median prices of different types of fuels reported by households in urban areas in a district are taken as prevailing prices in district.[26]

[25] IHDS data is publicly available from Inter-university Consortium for Political and Social Research (ICPSR). See http://ihds.info/ for more details.

[26] For rural sample, the IHDS data do not report price of coal, I impute price of coal for rural households using district level coal prices from NSS rural sample. District is the lowest level of geographical unit that can be identified in both

The women module of the IHDS was implemented to only those households that have a residing adult women in age 18-49. The women module contains questions that I use to assess women autonomy inside the household. I generate five variables to measure women bargaining power. The first one is the education gap which equals to the highest education level of female adults minus the highest education level of male adults in household. I create four additional indices that capture violence against women, involvement in decision making, freedom of movement, and financial independence of women. The indices are constructed by combining multiple questions available in the survey utilizing the principal components analysis (PCA). The PCA takes into account which measures are proxying the same concept as opposed to different concepts.

The first index, violence index, captures the magnitude of violence in the community. It is based on following questions: *In your community is it usual for husbands to beat their wives in each of the following situations?* 1) if she goes out without telling him; 2) if her natal family does not give expected money, jewelry or other items; 3) if she neglects the house or the children; 4) if she doesn't cook food properly; 5) if he suspects her of having relations with other men. The second index, *most say in decision-making*, captures the women's extent of say in various decision making. The index is based on *whether women have say in following decisions*: 1) what to cook on a daily basis; 2) whether to buy an expensive item such as a TV or fridge; 3) How many children to have; 4) what to do if a child falls sick; 5) to whom children should marry.[27] The third index, financial independence index, captures the magnitude of financial freedom enjoyed by women based on the following three questions asked to the respondent woman: 1) Do you yourself have any cash in hand to spend on household expenditures; 2) is your name on any bank account, 3) is your name on the ownership or rental papers for your home.[28] The fourth index, permission index, captures whether female in a household have less freedom to go out. It is constructed using the question: *Do you have to ask permission*

datasets for matching purposes.

[27]Nordman and Sharma (2016) create similar index by adding the binary responses.

[28]Nordman and Sharma (2016) also create a similar index by adding the binary responses in the three questions.

of your husband or a senior family member to go to ... 1) the local health center; 2) the home of relatives or friends in the village/neighborhood; 3) the Kirana shop.

In my data, there are total six fuels used for cooking—firewood, dung, crop residuals, coal/charcoal, kerosene, and LPG. IHDS questionnaire lists each fuel type and asks from the respondent whether the household has used the fuel for cooking purposes. The use of electricity as fuel type is not listed, however, according to 2011 Census data, only 0.10 percent of households in India listed electricity as their main cooking fuel (0.07 percent of rural households and 0.15 percent of urban households). Figure 6 and Figure 7 present the types of fuel used by rural and urban households, respectively. As evident from the figures, majority of the households in rural India use firewood, while majority of urban households use LPG. Following the literature, I treat firewood, dung, crop residuals, and coal/charcoal as dirty fuels, while kerosene and LPG are treated as clean fuels. Figure 8 presents the fuel stacking behavior in urban and rural households. As evident, fuel stacking is prevalent in both urban and rural areas, and importantly, the incidence of fuel stacking has increased between 2004 and 2011 in both urban and rural India.

Although, IHDS data also contain information on total expenditure on each type of fuel in the last 30 days, the expenditure information is available if the household bought the fuel from the market. For households who collected their own fuel, there is no available imputed value. This is problematic for traditional fuels as a large fraction of rural households use traditional fuels that are either collected from own land or other places. Moreover, since there is no quantity information on each type of fuel, it is not possible to impute expenditure for households who do not buy traditional fuels. Hence, I do not use information on fuel expenditure in my analysis.[29]

Table 10 provides the descriptive statistics of the variables used in my analysis. The explanatory variables used in the analysis include household head characteristics, household demographic characteristics including the social groups, household consumption expenditure as a proxy for income, fuel prices of alternative fuels at the village level (for

[29] Choumert-Nkolo et al. (2019) construct continuous indices using the share of expenditure of each type of fuel in total fuel expenditure.

rural households) or at the district level (for urban households), and different variables that capture the bargaining power of women as discussed earlier. For rural households, I also control for village infrastructure.

Indian society has historically been characterized by a high degree of social stratification governed by the caste system, which results in exclusion of certain groups from certain economic and social spheres. At the time of independence, the Indian Constitution identified the disadvantaged caste and tribes in a separate schedule of the constitution as Scheduled Castes and Scheduled Tribes (SC/STs), and extended affirmative action protection to these groups in the form of reserved seats in higher educational institutions, in public sector jobs, and in state legislatures as well as the Indian parliament. Other Backward Castes (OBCs) are a group of other backward castes grouped together, and the Government of India provided reserved positions for OBCs in public sector jobs in 1993. Muslims are the largest minority religious group in India, and according to the Government of India (2006), their performance on many economic and education indicators is comparable with that for SC/STs. I control for the social group by including indicators for belonging to Scheduled Castes (SC), Scheduled Tribes (ST), Other Backward Castes (OBC), or Muslim excluding non-Muslim others.

3. Empirical Framework

My analysis is based on an multinomial logit model with unobserved heterogeneity. Each household i faces j choices of cooking fuel at time t. Each fuel choice corresponds to certain level of utility. Household chooses the one for which the utility is the highest. In my setting, there are three choices at each time t: only dirty fuel ($j = 1$), a mix of clean and dirty fuels ($j = 2$), and only clean fuel ($j = 3$).[30]

Household i's indirect utility of a choice j at the time t in a random effects context can be specified as follows:

[30] In a study in Guatemala, Heltberg (2005) captures the stacking behavior of households by using three categories: only wood, only LPG, or LPG and charcoal for cooking.

$$V_{ijt} = X'_{it}\beta_j + \gamma_{ij} + \varepsilon_{ijt},\ t = 1, 2;\ j = 1, 2, 3 \tag{4}$$

where X_{it} is a matrix of observed explanatory variables that are expected to affect fuel choice and β_j is the choice specific parameter vector. γ_{ij} and ε_{ijt} are unobserved random components, where γ_{ij} is household and choice specific. At time t, household chooses the alternative with highest utility, V_{ijt}. Assuming that the ε_{ijt}'s are all independently distributed according to a type I extreme-value distribution, it can be shown that the conditional probability of household i choosing the category j at time t is given by:

$$P(Y_{it} = j \Big| X_{it}, \gamma_{i2}, \gamma_{i3}) = \frac{exp(X'_{it}\beta_j + \gamma_{ij})}{\sum_{c=1}^{3} exp(X'_{it}\beta_c + \gamma_{ic})} \tag{5}$$

where only dirty fuel is chosen as base outcome, so γ_{i1} and β_{i1} are normalized to zero.

If there is no unobserved heterogeneity across households: $\forall j : \gamma_{ij} = \gamma_j$, Eqn (2) will be a pooled multinomial logistic regression. As discussed in Alem et al. (2016), the standard multinomial logit model using the pooled sample (ignoring household heterogeneity) assumes that households' choices are independent, both within a choice (that is, for multiple observations across time of the same choice) and across all alternative choices made by the household over time. However, facing same set of choices in different time-periods, households are more likely to make similar choices over time. The random effects specification relaxes the assumption that multiple observations within a choice are independent. With this model, the choice probabilities for repeated choices made by household i share the same time-invariant unobserved heterogeneity γ_{ij}, where the household-specific effects act as a random variable that produces a correlation among the residuals for the same household within choices, but leaves the residuals independent across households. The multinomial random effects models are estimated using STATA gsem command.

4. Results

4.1. Household Fuel Choices

Table 11 presents the estimates of the panel multinomial logit modelsvwith random effects for rural and urban areas. The rural model includes additional village level characteristics compared to the urban model. Table 11 does not include variables capturing women bargaining power.[31] As evident from Table 11, the variances of the heterogeneity terms for both clean fuel and stacking are large and highly significant. Moreover, addition of random effects leads to a large increase in the log likelihood value over that for a standard pooled multinomial logit without random effects, and a likelihood ratio test rejects the pooled multinomial in favor of random effects. [32]

In Table 12, I present the average marginal effects from the random effects models reported in Table 11. Households belonging to the disadvantaged groups SCs, STs, and OBCs are more likely to use only dirty fuel in rural areas compared to households belonging to non-Muslim others group. Moreover, the disadvantaged category households are less likely to use only clean fuel or stack in rural areas compared to non-Muslim other households. In contrast to rural areas, SC/ST households in urban area are more likely to stack compared to non-Muslim other households. Importantly, similar to rural areas households belonging to disadvantaged categories (SC/STs and OBCs) are more likely to use only dirty fuel and less likely to rely on only clean fuel compared to households belonging to non-Muslim others group.

An increase in monthly per capita consumption increases the probability of use of both mixed and clean fuel in rural areas. Importantly, increase in consumption expenditure leads to a larger increase in the probability of mixed fuel use than increase in probability of only clean fuel in rural areas. As expected, increase in consumption expenditure reduces the probability of relying only on dirty fuel in rural areas. I find slightly different results

[31] In all of my estimations, I control for the six region. Ideally, I would like to control a finer geographical unit such as state (for urban sample) and districts (for rural sample). However, because of large sample size and presence of large number of indicators lead to convergence issues in the estimation.

[32] Fixed effect multinomial estimates are reported in appendix Table A6, however, I do not discuss the results here.

for urban households. While increase in consumption expenditure is associated with a significant decrease (increase) in the probably of use of only dirty (clean) fuel, increase in consumption expenditure reduces the use of mixed fuel marginally. It is noteworthy that there exists substantially large difference in the use of sole fuel between rural and urban areas. In 2011, only 69 (7) percent of rural households reported use of only dirty (clean) fuel, while 31 (23) percent urban households reported using only dirty (clean) fuel. Given the large differential in the use, the differences in rural-urban results regarding use of mixed fuel support the idea that households move from dirty fuel to mixed fuel and then to clean fuel. Many studies have studied the impact of income/consumption on fuel choice, and findings have been mixed. My findings in rural India are similar to Heltberg (2004, 2005) who use multiple country data and finds that with increase in income, households tend to add modern fuels to their mix as partial rather than perfect substitutes for traditional ones. Alem et al. (2016) suggest that households tend to switch to a multiple fuel–use strategy as their incomes rise, for reasons that include the reliability of supply and convenience of use of different stoves and fuel types. Hence, supply side considerations with differential initial use levels might be driving reasons behind differential impact of income on mixed fuel in rural areas compared to urban areas.

Increase in household size increases the probability of fuel stacking in both urban and rural areas, and reduces the probability of relying on only one type of fuel. Heltberg (2004) and Alem et al. (2016) also find that larger households are more likely to be involved in fuel stacking. A higher dependency ratio increases the probability of relying on only dirty fuel while reducing the probability of use of any clean fuel (either clean only or mixed). This is true both in urban and rural areas. Importantly, children are normally involved in collecting traditional fuel. More children reduce the cost of collecting traditional fuel; hence, increase the use of traditional dirty fuel. Households who derived their main income through salary or trade are less (more) likely to use dirty (clean) fuel in both urban and rural area compared to households whose main income source in cultivation/agriculture. Moreover, households whose main income source is salary or trade are more likely to use mixed fuel in rural areas, while less likely to use mixed fuel in urban areas.

Female-headed households in rural areas are less likely to use only dirty fuel and more likely to be involved in fuel stacking behavior. In urban areas, female headed households are more likely to rely only on clean fuel. It is important to note that the female-headed rural households are only marginally more likely to use clean fuel, while for the urban households the impact on the use of mixed fuel is insignificant. This suggests that females prefer to use clean fuel, however, because of supply constraints, they move to mixed fuel in rural areas, while they moves to clean fuel in urban areas. The existing literature also supports the idea that the female-headed households prefer modern fuels to traditional fuels (Farsi et al., 2007; Rao and Reddy, 2007; Rahut et al., 2014). This is generally attributed to the fact that women are often responsible for household cooking and thus are directly affected by the air pollution emitted from the burning of the dirty fuels. Age of head also reduces the use of only dirty fuel and increases the probability of household adopting clean fuel in both rural and urban areas. Muller and Yan (2018) suggest that this result implies clean fuels is more affordable for the elderly than the young people because the later facing liquidity constraints. The education of household head plays a role in fuel choices. More education is associated with a decline in the probability of relying on only dirty fuel, while increases the probability of relying only on clean fuel in both rural and urban areas. Consistent with other results, although more education increases the probability of fuel stacking in rural areas, it decreases the probability of fuel stacking in urban areas.

The prices of fuels have drawn considerable attention in fuel choice literature. I find that an increase in firewood prices reduces (increases) the probability of use of only dirty (clean) fuel in rural areas, and has no impact on the fuel stacking behavior. In urban areas, increase in firewood prices reduces probability of use of either only dirty fuel or mixed fuel, while increases the probability of use of only clean fuel. Thus, households tend to shift to clean fuel sources as price of firewood increases. Importantly, the magnitude of impacts of increase in firewood prices are much larger in urban areas compared to rural areas. This is possibly because, while the households in urban areas mostly rely on market for firewood, households in rural areas can self-collect firewood potentially reducing the impact of an increase in firewood prices. Increase in price of coal reduces the probability of using either clean or dirty fuel and increases the probability of use of mixed fuel in both urban and rural areas. It is noteworthy that less than 2 percent of rural households and close to 5 percent of urban households reported use coal for cooking. Increase in dung prices increases the probability of using only dirty fuel, while reduces the use of mixed fuel in rural areas. The estimates suggest that increasing dung prices also increases the probability of use of clean fuel; however, the magnitude of impact is marginal. Increase in use of dirty fuel with dung prices is counter-intuitive; however, as my indicator for dirty fuel contains four types of fuel, there is a possibility of cross-substitution. Some scholars suggest that cross-price effects can be an important driver of fuel substitution (Muller and Yan, 2018). For example, Peng et al. (2010) show that high coal prices increase the probability of choosing biomass in China, suggesting that coal and biomass may be substitutes.

Increase in kerosene prices decreases the use of only dirty fuel for rural households while increasing the probability of relying on only clean fuel or mixed fuels. For urban households increase in kerosene prices increases the use of only clean fuel but reduces the probability of use of dirty or mixed fuel. It is noteworthy that my clean fuel category includes both LPG and kerosene, and cross-substitution within clean fuel group may increase the use of clean fuel. Akpalu et al. (2011) find kerosene and LPG are substitutes. Gupta and Kohlin (2006) using data in India suggest a high cross-price elasticity of LPG with respect to Kerosene as well. Nonetheless, the decreased use of only dirty fuel as a result of increase in kerosene price is puzzling. However, similar results are reported in some other contexts. For example, Lay et al. (2013) report a statistically significant negative effect of the kerosene price on the choice of wood in Kenya. In fact, kerosene prices are not considered as an effective policy instrument to trigger fuel switching (Pitt, 1985; Lee, 2013; Akpalu et al., 2011). Interestingly, kerosene in India is one of the subsidized commodities for household use, and been distributed by the Public Distribution System (PDS) for decades. In other words, the actual kerosene prices faced by the households may be different than the market prices based on the households' eligibility for PDS and amount of PDS kerosene consumed. This introduces another uncertainty regarding the impact of market price for kerosene on clean fuel.

Increase in the price of LPG reduces use of mixed fuel and increases the use of dirty fuel for rural households, although the impact on dirty fuel is not statistically significant. Nonetheless, the impact of LPG price is much stronger in urban areas compared to rural areas, which is not surprising given a much higher incidence of LPG use in urban areas. For urban households, one percent increase in prices leads to 13.6 percentage points decline in use on clean fuel, and leads to a 14.3 percentage points increase in use of only dirty fuel. This is consistent with the evidence in the fuel choice literature that suggest a strong own-price effects for the demand of LPG (Farsi et al., 2007; Zhang and Kotani, 2012), and a substitution relationship between LPG and firewood (Heltberg, 2004; Sehjpal et al., 2014).

The availability of modern fuel like LPG should not be a big issue for urban households, however, it can be a key issue for rural households. Seventy eight percent of the urban households reported using LPG in 2012; however, only 30 percent of the rural households reported using LPG. Access to the LPG may be a contributing factor in the observed difference in the use of LPG across urban and rural areas besides other demand side factors. Increase in the distance to nearest town increases the use of dirty fuel and reduces the use of mixed fuel. There is no impact on the use of only clean fuel. Whether the village has paved road or not seems to an important determinant of fuel choice. Not having paved road lead to 4 percentage points increase in use of dirty fuel while it reduces the use of clean or mixed fuel. I also find evidence of social spillovers. Households residing in villages that reported clean fuel as their main source are more likely to use clean or mixed fuel and less likely to use dirty fuel. Households residing in villages that reported dirty fuel as their main cooking energy source are more likely to use dirty fuel and less likely to use clean fuel.

4.2. Women's Bargaining Power

Appendix Table A7 presents the results of multinomial random effects models that also include women bargaining power[33] , while Table 13 presents the average marginal impacts. Note that Table 13 is similar to Table 12 except that Table 13 models also include variables that capture the bargaining power of women. Since the bargaining information is constructed from the women module that was implemented only if the household has a residing adult women in age 18-49, inclusion of bargaining variables leads to loss of significant number of observations. However, the estimates of the rest of the variables remain qualitatively similar to what is presented in Table 12, and discussed earlier. Hence, I focus only on bargaining variables from the Table 13.

A greater education gap between female and male—namely greater women bargaining power— in the rural households results in a lower probability to use only dirty fuel but a higher probability to use only clean fuel and mixed fuel. The education gap makes urban households more likely to use only clean fuel, and less likely to use either dirty or mixed fuel. Similar findings are also reported in the Ahmad and Oliveira (2015) for urban India and Choumert-Nkolo et al. (2019) for Tanzania. Increase in women's financial independence increases the probability of use of clean fuel for urban households while reducing the probability of dirty or mixed fuels. For rural households also more financial independence lead to increase in probability of use of mixed fuel and reduction in probability of use of dirty fuel. The impact of financial independence on the use of clean fuel for rural households is marginal, however, positive. My result suggests that empowering women by education and individual finance could be effective policy instruments to accelerate the process of fuel switching.

[33]The results of multinomial logit fixed effect model with women bargaining power are reported in appendix Table A8, however, I do not discuss the results here.

Women greater bargaining power as captured by more say in household decisions has no significant impact on fuel choice for rural households. For urban households, however, more say in household decision marginally increases the probability of use of dirty fuel. A one standard deviation increase in say increases probability of use of dirty fuel by 0.6 percentage points for urban households. Violence index, that captures the perception of women about the amount of violence prevalent, increases the probability of using only dirty fuel or only clean fuel, while reduces the probability of using mixed fuel for rural households. For urban households, an increase in violence index increases the probability of use of only clean fuel while reducing use of mixed fuel. Recall that a larger violence index implies less bargaining power for women. Given the index is standardized, the impact of violence index remain marginal.[34] The permission index that captures the permission needed has no significant impact on the choice of fuel. Overall, the bargaining power of women that is associated with economic status (e.g. education or economic freedom) leads to increase in probability of household using clean fuel.

5. Conclusion

In this paper, I use two waves of India Human Development survey panel data to examine factors determining household fuel choice in rural and urban areas, separately. Majority of literature on fuel choice in India is based on cross-sectional data and thus does not allow for household heterogeneity. I contribute to the existing literature on India by using a panel data, and utilizing multinomial logit with random effects. Moreover, I also examine the impact of village infrastructure and women bargaining power on fuel choice.

I find considerable incidence of fuel stacking where households use both clean and dirty fuel together in both urban and rural areas. The use of only clean fuel remains low in rural areas, and between 2004 and 2011, there is an increase in fuel stacking not only in rural areas but also in urban areas. I find that the household characteristics that are positively associated with use on only clean fuel, such as education, per capita

[34] Moreover, the violence index captures respondent woman's perception about the status of women in the community, and may not be capturing the intra-household bargaining power of respondent women.

expenditure, are also positively associated with use of fuel stacking for rural households. For urban households, these characteristics have mostly no significant impact on fuel stacking. For rural households, I find access to paved road increases the probability of use of clean fuel suggesting supply side considerations play a role in adoption of clean fuel. I also find that female-headed households are more likely to adopt clean fuel, and increasing economic freedom of women is positively association with use of clean fuel.

CHAPTER III

IMPACTS OF CROSS-COUNTRY INFRASTRUCTURE ON BILATERAL TRADE: EVIDENCE FROM NEW SILK ROAD RAILWAYS

1. Introduction

Since implementing free market economic reforms in the 1970s, China has experienced a rapid economic growth. Growing economic power has enabled China to involve in more global economic policies and projects. Belt and Road Initiative (BRI) is an international trade and development strategy, proposed officially by President Xi, Jinping of China in 2013, to improve connectivity on a trans-continental scale. The "Belt" (Silk Road Economic Belt) links China to Central and South Asia and onward to Europe, while the "Road" (New Maritime Silk Road) links China to the nations of South East Asia, the Gulf Countries, North Africa, and on to Europe.

New Silk Road railways on the Belt have been being operated since 2011. These new railway services between China and Europe are cheaper than air freights and faster than ocean shipments. Trading times are particularly important for time sensitive products that are used as inputs in production processes, suggesting that reductions in shipping times are key in the presence of global value chains (Baniya, Rocha, and Ruta, 2019). The destination countries may face greater competition with China even though they have better access to Chinese market (Cosentino, et al., 2018) potentially affecting both imports (exports) from (to) China. Although Silk Road Railways in an ongoing project, as connection to more countries are introduced, I examine the impact of Silk Road railways on the bilateral trade of destination countries with China focusing on the destination countries that got railways connection in different years over 2011-2018. A major challenge in measuring causal effects of railways is that the selection of destination countries is not random but may depend on ancient Silk Road trading route, geopolitical goals, or China's longer-term economic strategy, etc. Moreover, the countries having had a tight trade relation with China may be more likely to participate in the project. To address the issue of endogenous placement of railways, I use the generalized synthetic control method (GSCM) proposed in Xu (2017). GSCM is quite suitable for time-series cross-sectional (TSCS) data and multiple treatment units receiving treatment at different times.

Although there is a growing literature on entire BRI, the literature focusing on Silk Road railways is limited.[35] Li, Bolton, and Westphal (2018) use Difference-in-Difference (DID) methodology and TSCS data from 28 countries over 2005-2014 to estimate the impact of Silk Road railways on bilateral trade. They find that although the railway connection does not affect the destination countries exports to China but destination countries imports from China increased by 29.1 percent post connection. The critical identifying assumption of DID is the "parallel trends," which states that in the absence of the treatment the average outcomes of treated and control units would have followed parallel paths. This assumption is not directly testable and generally supported by placebo that suggests parallel trend in pretreatment period. The presence of unobserved time-varying confounders causes the failure of this assumption.

Abadie, Diamond, and Hainmueller (2010) propose Synthetic Cohort Method (SCM) that matches both pretreatment covariates and outcomes between a treated unit and a set of control units and uses pretreatment periods as criteria for good matches. Specifically, it constructs a "synthetic control unit" as the counterfactual for the treated unit by reweighting the control units. However, it only applies to the case of one treated

[35]There exist literature that have examined the effect of BRI on income (e.g. Maliszewska and Van Der Mensbrugghe, 2019; Bird, Lebrand, and Venables, 2019), trade (e.g. Baniya, Rocha, and Ruta, 2019; Boffa, 2018; Zhang, et al, 2018), welfare (e.g. De Soyres, 2018), poverty and environment(e.g. Losos, 2019; Maliszewska and Van Der Mensbrugghe, 2019), or debt to GDP ratio (e.g. Bandiera and Tsiropoulos, 2019).

unit. Bai (2009) proposes an interactive fixed effects (IFE) strategy to model unobserved time-varying confounders semiparametrically. IFE strategy incorporates unit specific intercepts interacted with time-varying coefficients. The GSCM proposed in Xu (2017) unifies the synthetic control method with linear IFE model. It treats counterfactuals of treated units as missing data and makes out-of-sample predictions for posttreatment treated outcomes based on an IFE model (Xu, 2017). Moreover, it embeds a cross-validation scheme that selects the number of factors of the IFE model automatically. The GSCM addresses the lack of parallel trends between control and treated units, as it models the unobservable time-varying coefficients semi-parametrically by using the interactive fixed effects (IFE) model proposed in Bai (2009). In addition to that it corrects the bias of the IFE model when the treatment effects are heterogeneous across units (Maamoun, 2019).

This paper adds to the literature in following ways. First, the paper adds to limited literature on Silk Road Railways by using GSCM which provides a significant improvement over existing DID estimates as discussed above. Moreover, use of GSCM allows heterogeneous treatment effects across destination countries. Second, I use a longer period compared to Li, Bolton, and Westphal (2018) that looks only at the initial stage of connection. My sample include more destination countries and capture the impact over much longer time period.

My findings are following. I do not find significant impact on either destination countries exports to China or imports from China. Moreover, I find similar results when I look at different product category at one-digit ISO. I further divide destination countries based on whether they share boundaries with China or not. I find similar results. Although my results for exports are in accordance with results for exports presented in Li, Bolton, and Westphal (2018), my results that suggest insignificant impact for imports are in contrast with theirs. Their DID estimates based on data from 28 countries (7 Treated and 21 Controls) over 2005 - 2014 suggest an increase in exports post railway connection.[36] I discuss the differences my results with theirs in results section.[37]

Although my results suggest that provision of railway connection did not lead to significant change in destination countries exports (imports) to (from) China, I cannot rule out substitution of volume between different modes of transportation. Since my analysis is based on aggregated trade reported for the countries not based on transportation method, I cannot rule out substitution across alternative modes of transportations.

This paper is organized as follows. Section 2 provide a brief introduction of Belt Road Railways, Section 3 describes the data. Section 4 details the empirical strategy and Section 5 presents the results. Section 6 concludes.

2. Background: New Silk Road Railways

Belt and Road Initiative is a long-term and transcontinental investment and policy program, which unveiled by China‘s president Xi Jinping in 2013. China has invested exceeded 100 billion U.S. dollars in the program. The purpose of BRI is to accelerate regional integration through investment and infrastructure development. It aims at enhancing the connectivity of Asian, European and African continents and their adjacent seas along Silk Road Economic Belt (Belt)[38] and New Maritime Silk Road (Road). Currently, there are at least 71 participating countries.

Figure 9 shows the overall project routes: the Belt, a long-term plan of infrastructural development, has six major land corridors (brown and orange lines), while the Road includes one maritime corridor (the blue line). New Silk Road railways belongs to projects of the Belt. My research interest is at the freight rail services primarily along New Eurasia Land Bridge Economic Corridor and China-Mongolia-Russia Economic Corridor. Until 2018, there are at least nine rail routes belongs to these two corridors.

Along New Eurasia Land Bridge Economic Corridor, the freight rails transfer prod-

[36] In their paper, they do not specify the number of destination countries. I assume there are 7 destination countries based on the time line I believe.

[37] In essence, I replicated the DID results using the similar data as used in Li, Bolton, and Westphal (2018). I find insignificant impacts on both exports and imports using the same sample.

[38] I called the “Belt” in the rest of content

ucts made in China mainly run through Dzungarian Gate to Central Asia and Europe. For example, the operation of the first New silk Road railway, Yuxinou, is along this corridor. In 2011, before official BRI announcement, Yuxinou started operation from Chongqing, China, passed through Kazakhstan, Russia, Belarus, Poland, and arrived in Duisburg, Germany. The railways operating along China-Mongolia-Russia Economic Corridor mainly pass Mongolia and Russia through Manzhouli. Dzungarian Gate and Manzhouli are the important land ports for the two corridors.

The railways of my interest mainly connect China with countries in Central Asia and Europe, so I am going to see if the railways indeed have impacts on the destination countries comparing to other Central Asian and European countries. The railways stop by 15 Central Asian and European countries including Mongolia. In addition to the countries along the first rail route, the railways also arrived Czech Republic in 2012; France, Mongolia, and Uzbekistan in 2014; Iran, Spain, and Tajikistan in 2015; Netherlands in 2016; United Kingdom and Hungary in 2018.[39]

3. Data

I assemble data set from various sources to construct a panel dataset for 65 European and Central Asian countries.[40] The bilateral trade data is sourced from World Integrated Trade Solution (WITS). In addition to using aggregated bilateral trade data, I also use bilateral trade by Standard International Trade Classification (SITC) Revision 3.[41] The characteristics of countries such as export-to-GDP ratio, import-to-GDP ratio, GDP, labor force participation, total population, average tariff, and exchange rate are taken from World Development Indicators (WDI).[42]

For the connection of railways data, I define a country receiving railway connection by

[39] Too many missing values in the trade outcomes for Uzbekistan and Tajikistan, so these two countries are excluded in my analysis.

[40] I exclude island countries except Britain which has been connected through Silk Road railways.

[41] I select SITC3 because trade data under SITC3 have less missing values.

[42] For few countries' missing exchange rate data in WDI are substituted with data from Federal Reserve Economic Data (FRED) and Eurostat. I replace the missing tariff data in WDI using previous or earlier years' values since tariff may not change in a few years.

a given year when the first rail fright arrived between 2011 - 2018.[43] The main outcome variables are log of gross import from China and log of gross export to China.[44] My main sample include 59 countries and 13 of which are destination countries.[45]

Figure 10 shows the treatment and control countries over year.[46] 13 countries connect with China by railways at different time period. The rest of the countries serve as control countries. Table 14 shows summary statistics of treatment and control groups. On average, treatment group has larger gross import and gross export and a lower tariff rate than control group.

4. Empirical Strategy

First, I estimate the traditional DID model by estimating the following equation:

$$Y_{it} = \delta Connected_{it} + X_{it}'\beta + \theta_i + \eta_t + \epsilon_{it} \tag{6}$$

where the dependent variable Y_{it} is the logarithm of the gross import (or gross export) for country i from (to) China in year t. $Connected_{it}$ is a binary indicator that takes a value one after country i become a destination countries in year t. X_{it} is a matrix of time-varying country characteristics: log of GDP, log of population, labor force participation rate, log of exchange rate, average tariff rate, share of exports, imports, agriculture, and manufacturing in GDP. θ_i and η_t are country and year fixed effects to capture unobserved country-specific factor and time-variant shocks.[47] ϵ_{it} is the idiosyncratic error term. The parameter of interest, δ, measures the causal effect of railways on trade outcomes under 'parallel' trend assumption. Standard errors are clustered by country. As discussed in introduction section, because of time-varying unobserbales, the parallel trend assumption

[43] There is no official guide book of New Silk Roads railways, so I refer relevant News, Baidu, Wikipedia, and Li, Bolton, and Westphal (2018).

[44] Gross import/export is the sum of import/export and re-import/re-export.

[45] I drop 6 countries with too many missing values in outcome or control variables: Uzbekistan, Djibouti, Iraq, Libya, Turkmenistan, and Andorra.

[46] The data is an unbalanced panel with a few missing values represented by white cells.

[47] I do not control for the time-invariant country specific variables generally used in gravity models as they will be absorbed by country or time fixed effect. For example, the partner country (China) GDP is absorbed by year fixed effects. Similarly, country specific variables such as sharing the same official language or border or past colonial relationship, the strength of market penetration are absorbed by country fixed effects.

may not hold. To allow for the possibility of parallel trend assumption not holding, I estimate Generalized Synthetic Control Analysis (GSCM) proposed in Xu (2017).

4.1. Generalized Synthetic Control Analysis

Following the framework and notations in Xu (2017), let τ and C denote the sets of treatment and control units, respectively. N_τ and N_C are the numbers of treatment and control units, so the total number of units, is $N = N_\tau + N_C$. All units are observed for T periods. Assume T_0 be the number of pre-treatment periods. The outcome variable Y_{it} can be expressed in the following functional form:

$$Y_{it} = \delta_{it} Connected_{it} + x'_{it}\beta + \lambda'_i f_t + \epsilon_{it} \tag{7}$$

where δ_{it} is the heterogeneous treatment effect for country i in year t. f_t is an $(r \times 1)$ vector of unobserved common factors while λ_i is an $(r \times 1)$ vector of unknown factor loadings representing country specific intercepts. ϵ_{it} represents unobserved idiosyncratic shocks for country i in year t.

Based on the aforementioned functional form, $Y_{it}(1)$ and $Y_{it}(0)$, the potential outcomes for country i in treatment years or in pre-treatment years, respectively, can be expressed by the following equations:

$$Y_{it}(1) = \delta_{it} + x'_{it}\beta + \lambda'_i f_t + \epsilon_{it} \tag{8}$$

$$Y_{it}(0) = x'_{it}\beta + \lambda'_i f_t + \epsilon_{it}. \tag{9}$$

This treatment effect is captured by δ_{it}

$$\delta_{it} = Y_{it}(1) - Y_{it}(0), t > T_0 \tag{10}$$

Stacking all control units together, I have:

$$Y_{co} = X_{co}\beta + F\Lambda'_{co} + \epsilon_{co} \tag{11}$$

where $Y_{co} = [Y_1, Y_2, ..., Y_{N_{co}}]$ and $\epsilon_{co} = [\epsilon_1, \epsilon_2, ..., \epsilon_{N_{co}}]$ are $T \times N_{co}$ matrices. X_{co} is a three-dimensional $(T \times N_{co} \times p)$ matrix, and $\Lambda_{co} == [\lambda_1, \lambda_2, ..., \lambda_{N_{co}}]'$ is a $(N_{co} \times r)$ matrix, $X'_{co}\beta$ and $F\Lambda'_C$ are also $T \times N_{co}$ matrices. As discussed in Xu (2017), to identify β, F, Λ_{co}, two sets of constraints on the factors and factor loadings are needed: all factor are normalized, and they are orthogonal to each other, i.e.: $F'F/T = I_r$ and $\Lambda'_{co}\Lambda_{co}$ =diagonal. The number of factor, r, is chosen by a cross-validation procedure as described in Xu (2017). The main aim of the GSCM is to find the effect of the treatment by finding the average difference between the treated unit(s) and its counterfactual. The average treatment effect on the treated (ATT) in the year t when $t > T_0$ can be given by following equation:

$$ATT_{t,t>T_0} = \frac{1}{N_{tr}} \sum_{i \in \tau} [Y_{it}(1) - Y_{it}(0)] = \frac{1}{N_{tr}} \sum_{i \in \tau} \delta_{it}. \tag{12}$$

$Y_{it}(1)$ is observable for treated countries in post-treatment periods, but $Y_{it}(0)$ is missing. To construct treated counterfactuals for each treated countries in the post-treatment period, I first have to estimate $\hat{\beta}$, $\hat{F}$, and $\hat{\Lambda}_{co}$:

$$(\hat{\beta}, \hat{F}, \hat{\Lambda}_{co}) = \underset{\tilde{\beta}, \tilde{F}, \tilde{\Lambda}_{co}}{\operatorname{argmin}} \sum_{i \in C} (Y_i - X_i\tilde{\beta} - \tilde{F}\tilde{\Lambda}_i)'(Y_i - X_i\tilde{\beta} - \tilde{F}\tilde{\Lambda}_i)$$

$$s.t. \qquad \tilde{F}'\tilde{F}/T = I_r \quad and \quad \tilde{\Lambda_C}'\tilde{\Lambda_C} = diagonal.$$

The second step estimates factor loadings for each treated country by minimizing the mean squared error of the predicted treated outcome in the pre-treatment periods:

$$\begin{aligned} \hat{\lambda}_i &= \underset{\hat{\lambda}_i}{\operatorname{argmin}} (Y_i^0 - X_i^0\hat{\beta} - \hat{F}^0\tilde{\lambda}_i)'(Y_i^0 - X_i^0\hat{\beta} - \hat{F}^0\tilde{\lambda}_i) \\ &= (\hat{F^0}'\hat{F^0})^{-1}\hat{F^0}'(Y_i^0 - X_i^0\hat{\beta}), \quad i \in \tau, \end{aligned}$$

where $\hat{\beta}$ and $\hat{F^0}$ are estimated by the first step, and the superscripts "0"s denote the pre-treatment periods. The third step uses the estimation, $\hat{\beta}$, $\hat{F}$, and $\hat{\Lambda}_{co}$, to calculate treated counterfactuals:

$$\hat{Y}_{it}(0) = x_{it}'\hat{\beta} + \hat{\lambda}_i'\hat{f}_t, i \in \tau, \quad t < T_0.$$

Hence, the ATT estimator can be obtained: $\widehat{ATT}_t = (1/N_{tr})\sum_{i\in\tau}[Y_{it}(1) - \hat{Y}_{it}(0)]$ for $t > T_0$.

5. Results

5.1. Estimates of Difference-in-Difference Model

Column (1) of Table 15, presents the DID estimate for gross import when covariates are not controlled for. Although the DID estimate suggest positive impact of railways on destination countries imports from China, I cannot rule out no impact. Importantly, when I control for country time-varying characteristics in column (2) of Table 15,, the sign of DID estimate flips, though the DID estimate remains statistically insignificant. In column (3) of Table 15, I also control for country-specific linear time trends that allow each country to follow differential trend. The sign of DID flips positive again. Although the DID estimate remains insignificant in all specifications, the estimate switching signs create uncertainty about the direction of impact if any.

In column(1)-(3) of Table 16, I present similar DID results for gross exports of destination countries to China. The DID estimate without controls in column (1) presents a negative but insignificant impact of the railways on gross exports. The DID estimate in column (2) of Table 16 that controls for country characteristics suggest that the railway connection lead to a significant decline in destination countries gross exports to China. However, when I allow for differential trends in column (3) of Table 16, the DID estimate turns positive, although statistically significant.

Thus, based on DID estimates, I conclude that the connection of silk railway did not lead to a significant impact of destination countries imports from China. In terms of exports to China also, I do not find conclusive evidence of impact. These results are in contrast to the DID results presented in Li, Bolton, and Westphal (2018). They present the DID specifications which is similar to my column (2) specification that imposes similar trend across countries. They use a panel data from 28 countries over 2005-2014 that include 7 treated countries. They report no significant impact on destination counties exports to china, however, a significant increase in imports from China. As discussed earlier following a similar specification reported in column (2) of Table 15 and Table 16, I do not find statistically significant impact on imports, while a negative impact on exports. To reconcile the differences in DID estimates, I re-estimate restricting my sample similar to theirs. However, I could not replicate their results and find results which are similar to my full sample results.[48]

5.2. Estimates of Generalized Synthetic Control Method

The two-way fixed effects model referred as DID presented in earlier section assumes a constant treatment effect both across countries and over time. Next, I apply the GSCM to examine how the railways affect bilateral trade. The results of GSC estimation on gross import and gross export are presented in the column (4) and (5) of Table 15 and Table 16, respectively. First, I look into the effect of railways on gross import through a simple GSC model without controlling for country characteristics. In column (4), no characteristics are controlled for, while in column (5) characteristics are controlled for. Similar to DID, both these specifications impose additive state and year fixed effects. In the column (4) of Table 15, the GSCM estimate suggest that railway connection led to decline in imports from China. Controlling for characteristics in column (5) strengthens the magnitude of the impact, however, the GSCM estimate remain statistical insignificant at conventional level (p-value=0.058). The cross validation scheme in column (5) finds one unobserved factor to be important and after conditioning on the unobserved factor and additive fixed effects, the estimated ATT based on the GSCM suggests a 19 percent reduction in imports in the BRI countries as compared to the counterfactual that represents the

[48] The replicated results are available from the corresponding author. It is worth pointing out that in Table 3 and Table 5 of Li, Bolton, and Westphal (2018), the authors have reported controlling country and year fixed effects. However, they also report coefficient for country-specific time-invariant variables such as whether country is landlocked or not and distance, and year-specific variable e.g. Chinese GDP.

would-have-been scenario in the absence of the railway connection.[49]

In Figure 11a, I plot the actual import data of destination countries against the counterfactual generated through GSCM. As observed, the counterfactual is indistinguishable from treatment group in pre-treatment period suggesting that treated group should be exhibiting similar pattern as counterfactual in post-treatment period in the absence of treatment. Figure 11b demonstrate the impact of railway connection up to 5 periods after connection. Although the estimates suggest a negative impact on imports, I cannot rule out no impact as the confidence intervals include zero. Similarly, I do not find statistically significant impact on exports of destination countries (column (4) of Table 16). In Figure 12a, I plot the export of treated group and counterfactual to demonstrate that GSCM counterfactual mimic the treated group in pre-treatment period. Figure 12b plots the impact of railway up to 5 periods post-treatment, and suggest no significant impact on exports.

To check robustness of baseline GSC estimates, I also employ the matrix completion (MC) method proposed in Athey et al. (2017). Distinct from GSCM that using uses information from control units, MC method utilizes both control and treatment group information in the pre-treatment period to construct the counterfactuals of treated unit. The results for gross import and gross export are shown in the Appendix figures, Figure A7 and Figure A8. I do not find significant impact of railway connection on destination countries import from China. However, the MC method suggests negative impact on exports after couple of periods post-treatment.

5.3. The Effect of New Silk Road Railways on Trade by Classification of the Commodities and Destination Countries

In this section, I first explore heterogeneity in the effect of railways on trade by distinct product categories. For example, China is the world's largest manufacturer, so the railways may help to export its manufacturing products. I examine the outcomes for six

[49] Based on Xu (2017) the unobservable factors found are not directly interpretable. The one factor in my case may simply refer to a time trend that has a heterogeneous effect on the different countries (Maamoun, 2019).

product categories as a result of too many missing values in other categories of outcomes. Figure A9 and Figure A10 shows the effect of railway connection on gross import and export of the following product categories: food and live animals (SITC 0), crude materials (SITC 2), chemicals and related products (SITC 5), manufactured good (SITC 6), machinery and transport equipment (SITC 7), and miscellaneous manufactured articles (SITC 8). I do not find significant impact of railways on exports and imports of different product groups.

I further explore impact of railways based on dividing treated countries in two groups: countries that share boundary with China vs countries that do not share boundary with China. The results are shown in Figure A11 to Figure A14. I do not find impacts to vary based on location of treatment countries.[50]

6. Conclusion

New Silk Road railways, one of the projects under Belt and Road Initiatives, have been operational since 2011 and connect China to countries in Europe and Central Asia. Over time more countries got connected through railways. I examine, the impact of silk road railways on destination countries exports (and imports) to (from) China using generalized synthetic control method (GSCM). Given the time-series cross-section nature of data and different countries receiving treatment on different years, GSCM is suitable methodology that account for endogenous placement of treatment.

I do not find significant impact of railway connection on either destination countries exports to China or destination countries imports from China. My findings of no significant impacts on bilateral trades are striking given the amount of investment in Silk Road railways. However, there are many caveats with my study. First, given the trade data availability till 2018 and many countries having received railways only few years before 2018, my estimates capture only impacts for few years post railway connection. Given the long-term nature of investment, it may take time for effects to be material-

[50] I also examine the effect of railway connection on each destination countries, which are plotted in Figure A15 and Figure A16. Again, I do not find significant impact on each treated country bilateral trade with China.

ized. Second, since trade data for countries are not reported by means of transportation, I cannot rule out substitution across different modes of transportation. Moreover, the control countries may participate in the BRI project through New Maritime Silk Road reducing the impact of Silk Road railway on treated countries.

REFERENCES

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies 72*(1), 1–19.

Ahmad, S. and J. A. P. de Oliveira (2015). Fuel switching in slum and non-slum households in urban India. *Journal of Cleaner Production 94*, 130–136. ISBN: 0959-6526 Publisher: Elsevier.

Akpalu, W., I. Dasmani, and P. B. Aglobitse (2011). Demand for cooking fuels in a developing country: To what extent do taste and preferences matter? *Energy Policy 39*(10), 6525–6531. ISBN: 0301-4215 Publisher: Elsevier.

Alem, Y., A. D. Beyene, G. Köhlin, and A. Mekonnen (2016). Modeling household cooking fuel choice: A panel multinomial logit approach. *Energy Economics 59*, 129–137. Publisher: Elsevier.

Ananat, E. O., A. Gassman-Pines, and C. Gibson-Davis (2013, December). Community-Wide Job Loss and Teenage Fertility: Evidence From North Carolina. *Demography 50*(6), 2151–2171.

Andrews, D. W. (1993). Tests for parameter instability and structural change with unknown change poin. *Econometrica: Journal of the Econometric Society*, 821–856.

Ariel Kalil and Rebecca M. Ryan (2010). Mothers' Economic Conditions and Sources of Support in Fragile Families. *The Future of Children 20*(2), 39–61.

Athey, S., M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi (2018). Matrix completion methods for causal panel data models. Technical report, National Bureau of Economic Research.

Autor, D., D. Dorn, and G. Hanson (2019, September). When Work Disappears: Manufacturing Decline and the Falling Marriage Market Value of Young Men. *American Economic Review: Insights 1*(2), 161–178.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica 77*(4), 1229–1279. ISBN: 0012-9682 Publisher: Wiley Online Library.

Bandiera, L. and V. Tsiropoulos (2019, June). *A Framework to Assess Debt Sustainability and Fiscal Risks under the Belt and Road Initiative.* Policy Research Working Papers. The World Bank.

Baniya, S., N. Rocha, and M. Ruta (2019, January). *Trade Effects of the New Silk Road: A Gravity Analysis.* Policy Research Working Papers. The World Bank.

Becker, G. S. (1960). An economic analysis of fertility. In *Demographic and economic change in developed countries*, pp. 209–240. Columbia University Press.

Becker, G. S. (1965). A Theory of the Allocation of Time. *The economic journal*, 493–517.

Becker, G. S. (1973, July). A Theory of Marriage: Part I. *Journal of Political Economy 81*(4), 813–846.

Becker, G. S. (1974, March). A Theory of Marriage: Part II. *Journal of Political Economy 82*(2, Part 2), S11–S26.

Becker, G. S. (1981). A Treatise on the Family. *NBER Books*.

Bird, J. H., M. S. M. Lebrand, and A. J. Venables (2019). *The Belt and Road Initiative: Reshaping Economic Geography in Central Asia?* The World Bank.

Bitler, M. and H. Hoynes (2015, May). Living Arrangements, Doubling Up, and the Great Recession: Was This Time Different? *American Economic Review 105*(5), 166–170.

Black, D. A., T. G. McKinnish, and S. G. Sanders (2003). Does the availability of high-wage jobs for low-skilled men affect welfare expenditures? Evidence from shocks to the steel and coal industries. *Journal of Public Economics 87*(9-10), 1921–1942.

Blossfeld, H.-P., E. Klijzing, M. Mills, and K. Kurz (2006). *Globalization, Uncertainty and Youth in Society: The Losers in a Globalizing World*. Routledge.

Boffa, M. (2018, May). *Trade Linkages Between the Belt and Road Economies*. Policy Research Working Papers. The World Bank.

Bowmaker, S. W. and P. M. Emerson (2015, June). Bricks, Mortar, and Wedding Bells: Does the Cost of Housing Affect the Marriage Rate in the US? *Eastern Economic Journal 41*(3), 411–429.

Bruce, N., R. Perez-Padilla, and R. Albalak (2000). Indoor air pollution in developing countries: a major environmental and public health challenge. *Bulletin of the World Health organization 78*, 1078–1092. ISBN: 0042-9686 Publisher: SciELO Public Health.

Burke, P. J. and G. Dundas (2015). Female labor force participation and household dependence on biomass energy: evidence from national longitudinal data. *World Development 67*, 424–437. ISBN: 0305-750X Publisher: Elsevier.

Campbell, J. Y. and J. F. Cocco (2007). How do house prices affect consumption? Evidence from micro data. *Journal of monetary Economics 54*(3), 591–621.

Chan, M. H. T. (2018, July). The Belt and Road Initiative – the New Silk Road: a research agenda. *Journal of Contemporary East Asia Studies 7*(2), 104–123.

Charles, K. K., E. Hurst, and M. J. Notowidigdo (2018, October). Housing Booms and Busts, Labor Market Opportunities, and College Attendance. *American Economic Review 108*(10), 2947–2994.

Cheng, C.-y. and J. Urpelainen (2014). Fuel stacking in India: Changes in the cooking and lighting mix, 1987–2010. *Energy 76*, 306–317. ISBN: 0360-5442 Publisher: Elsevier.

Chetty, R., L. Sándor, and A. Szeidl (2017, June). The Effect of Housing on Portfolio Choice: The Effect of Housing on Portfolio Choice. *The Journal of Finance 72*(3), 1171–1212.

Choumert-Nkolo, J., P. C. Motel, and L. Le Roux (2019). Stacking up the ladder: A panel data analysis of Tanzanian household energy choices. *World Development 115*, 222–235. ISBN: 0305-750X Publisher: Elsevier.

Clark, W. A. (2012, July). Do women delay family formation in expensive housing markets? *Demographic Research 27*, 1–24.

Coile, C. C. and P. B. Levine (2011). The market crash and mass layoffs: How the current economic crisis may affect retirement. *The BE Journal of Economic Analysis & Policy 11*(1).

Corradin, S. and A. Popov (2015). House prices, home equity borrowing, and entrepreneurship. *The Review of Financial Studies 28*(8), 2399–2428.

Cosentino, B., D. Dunmore, S. Ellis, A. Preti, D. Ranghetti, and C. Routaboul (2018). Research for TRAN Committee: The new Silk Route–opportunities and challenges for EU transport. *DIRECTORATE-GENERAL FOR INTERNAL POLICIES. Policy Department for Structural and Cohesion Policies. Transport and Tourism.*

Das, S., H. De Groote, and B. Behera (2014). Determinants of household energy use in Bhutan. *Energy 69*, 661–672. ISBN: 0360-5442 Publisher: Elsevier.

De Soyres, F. M. M. R., A. Mulabdic, and M. Ruta (2019). *Common transport infrastructure: A quantitative model and estimates from the Belt and Road Initiative.* The World Bank.

Desai, S. and R. Vanneman (2009). National Council of Applied Economic Research.(2005),"India Human Development Survey (IHDS)," Computer file, ICPSR22626-v5. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor] 622.*

Desai, S. and R. Vanneman (2015). India human development survey-ii (ihds-ii), 2011-12. *ICPSR36151-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor] 31.*

Dettling, L. J. and M. S. Kearney (2014, February). House prices and birth rates: The impact of the real estate market on the decision to have a baby. *Journal of Public Economics 110*, 82–100.

Dietz, R. D. and D. R. Haurin (2003, November). The social and private micro-level consequences of homeownership. *Journal of Urban Economics 54*(3), 401–450.

Engelhardt, G. V. (2003). Nominal loss aversion, housing equity constraints, and household mobility: evidence from the United States. *Journal of urban Economics 53*(1), 171–195.

Fardella, E. and G. Prodi (2017, September). The Belt and Road Initiative Impact on Europe: An Italian Perspective. *China & World Economy 25*(5), 125–138.

Farnham, M., L. Schmidt, and P. Sevak (2011, May). House Prices and Marital Stability. *American Economic Review 101*(3), 615–619.

Farnham, M. and P. Sevak (2015). Housing wealth and retirement timing. *CESifo Economic Studies 62*(1), 26–46.

Farsi, M., M. Filippini, and S. Pachauri (2007). Fuel choices in urban Indian households. *Environment and Development Economics 12*(6), 757–774. ISBN: 1469-4395 Publisher: Cambridge University Press.

Ferreira, F. and J. Gyourko (2011, August). Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009. NBER Working Paper 17374, National Bureau of Economic Research, Cambridge, MA.

Ferreira, F., J. Gyourko, and J. Tracy (2010). Housing busts and household mobility. *Journal of urban Economics 68*(1), 34–45.

Gangopadhyay, S., B. Ramaswami, and W. Wadhwa (2003). Access of the poor to modern household fuels in India. *SERFA Report for the World Bank*.

Gangopadhyay, S., B. Ramaswami, and W. Wadhwa (2005). Reducing subsidies on household fuels in India: how will it affect the poor? *Energy Policy 33*(18), 2326–2336. ISBN: 0301-4215 Publisher: Elsevier.

Ghent, A. C. and M. T. Owyang (2010). Is housing the business cycle? Evidence from US cities. *Journal of Urban Economics 67*(3), 336–351.

Gibson-Davis, C. M. (2009). Money, marriage, and children: Testing the financial expectations and family formation theory. *Journal of Marriage and Family 71*(1), 146–160.

Gibson-Davis, C. M., K. Edin, and S. McLanahan (2005). High hopes but even higher expectations: The retreat from marriage among low-income couples. *Journal of marriage and family 67*(5), 1301–1312.

Gupta, G. and G. Köhlin (2006). Preferences for domestic fuel: analysis with socio-economic factors and rankings in Kolkata, India. *Ecological Economics 57*(1), 107–121. ISBN: 0921-8009 Publisher: Elsevier.

Heltberg, R. (2004). Fuel switching: evidence from eight developing countries. *Energy economics 26*(5), 869–887. ISBN: 0140-9883 Publisher: Elsevier.

Heltberg, R. (2005). Factors determining household fuel choice in Guatemala. *Environment and development economics 10*(3), 337–361. ISBN: 1469-4395 Publisher: Cambridge University Press.

Hillman, J. E. (2018). China's Belt and Road is Full of Holes. *Center for Strategic International Studies*.

Hosier, R. H. and J. Dowd (1987). Household fuel choice in Zimbabwe: an empirical test of the energy ladder hypothesis. *Resources and energy 9*(4), 347–361. ISBN: 0165-0572 Publisher: Elsevier.

Houthakker, H. S. and L. D. Taylor (1970). Consumer demand in the United States. *Cambridge, MA: Harvard University Press*.

Jialing, Z., L. Chunla, Y. Guoqing, and T. Zhipeng (2015). Spatial patterns and economic effects of China's trade with countries along the Belt and Road. *Progress in Geography 34*(5), 598–605.

Jingchao, Z. and K. Kotani (2012). The determinants of household energy demand in rural Beijing: Can environmentally friendly technologies be effective? *Energy Economics 34*(2), 381–388. ISBN: 0140-9883 Publisher: Elsevier.

Jones, L. E., A. Schoonbroodt, and M. Tertilt (2008). Fertility theories: can they explain the negative fertility-income relationship? NBER Working Paper 14266, National Bureau of Economic Research.

Kalil, A. and R. M. Ryan (2010). Mothers' economic conditions and sources of support in fragile families. *The Future of Children 20*(2), 39–61.

Kaygusuz, K. (2011). Energy services and energy poverty for sustainable rural development. *Renewable and sustainable energy reviews 15*(2), 936–947. ISBN: 1364-0321 Publisher: Elsevier.

Kearney, M. S. and P. B. Levine (2012, May). Why is the Teen Birth Rate in the United States So High and Why Does It Matter? *Journal of Economic Perspectives 26*(2), 141–166.

Kearney, M. S. and R. Wilson (2018, October). Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence from the Fracking Boom. *The Review of Economics and Statistics 100*(4), 678–690.

Klarl, T. (2016). The nexus between housing and GDP re-visited: A wavelet coherence view on housing and GDP for the US. *Econ Bull 36*(2), 704–720.

Klein, J. (2017, December). House Price Shocks and Individual Divorce Risk in the United States. *Journal of Family and Economic Issues 38*(4), 628–649.

Laeven, L. and A. Popov (2016). A lost generation? education decisions and employment outcomes during the us housing boom-bust cycle of the 2000s. *American Economic Review 106*(5), 630–35.

Lafortune, J. and C. Low (2017). Betting the House: How Assets Influence Marriage Selection, Marital Stability, and Child Investments. IZA Discussion Paper 11176, IZA.

Lay, J., J. Ondraczek, and J. Stoever (2013). Renewables in the energy transition: Evidence on solar home systems and lighting fuel choice in Kenya. *Energy Economics 40*, 350–359. ISBN: 0140-9883 Publisher: Elsevier.

Leach, G. (1992). The energy transition. *Energy policy 20*(2), 116–123. ISBN: 0301-4215 Publisher: Elsevier.

Leach, G. and R. Mearns (2013). *Beyond the woodfuel crisis: people, land and trees in Africa*. Routledge.

Leamer, E. E. (2007). Housing is the business cycle. NBER Working Paper 13428, National Bureau of Economic Research.

Lee, L. Y.-T. (2013). Household energy mix in Uganda. *Energy Economics 39*, 252–261. ISBN: 0140-9883 Publisher: Elsevier.

Li, Y., K. Bolton, and T. Westphal (2018, July). The effect of the New Silk Road railways on aggregate trade volumes between China and Europe. *Journal of Chinese Economic and Business Studies 16*(3), 275–292.

Lindo, J. M. (2010). Are children really inferior goods? Evidence from displacement-driven income shocks. *Journal of Human Resources 45*(2), 301–327.

Losos, E. C., A. Pfaff, L. P. Olander, S. Mason, and S. Morgan (2019, January). *Reducing Environmental Risks from Belt and Road Initiative Investments in Transportation Infrastructure*. Policy Research Working Papers. The World Bank.

Lovenheim, M. F. and K. J. Mumford (2013, May). Do Family Wealth Shocks Affect Fertility Choices? Evidence from the Housing Market. *Review of Economics and Statistics 95*(2), 464–475.

Maamoun, N. (2019). The Kyoto protocol: Empirical evidence of a hidden success. *Journal of Environmental Economics and Management 95*, 227–256. ISBN: 0095-0696 Publisher: Elsevier.

Maliszewska, M. and D. Van Der Mensbrugghe (2019, April). *The Belt and Road Initiative: Economic, Poverty and Environmental Impacts*. Policy Research Working Papers. The World Bank.

Masera, O. R., B. D. Saatkamp, and D. M. Kammen (2000). From linear fuel switching to multiple cooking strategies: a critique and alternative to the energy ladder model. *World development 28*(12), 2083–2103. ISBN: 0305-750X Publisher: Elsevier.

Mayer, C. (2011). Housing bubbles: A survey. *Annu. Rev. Econ. 3*(1), 559–577.

Mekonnen, A. and G. Köhlin (2009). Determinants of household fuel choice in major cities in Ethiopia. Working Papers in Economics 399, University of Gotenburg, Sweden. ISBN: 1403-2465.

Mian, A., K. Rao, and A. Sufi (2013). Household balance sheets, consumption, and the economic slump. *The Quarterly Journal of Economics 128*(4), 1687–1726.

Mian, A. and A. Sufi (2014, November). What Explains the 2007-2009 Drop in Employment?: The 2007-2009 Drop in Employment. *Econometrica 82*(6), 2197–2223.

Mincer, J. (1962). Labor force participation of married women: A study of labor supply. In *Aspects of labor economics*, pp. 63–105. Princeton University Press.

Muller, C. and H. Yan (2018). Household fuel use in developing countries: Review of theory and evidence. *Energy Economics 70*, 429–439. ISBN: 0140-9883 Publisher: Elsevier.

Nordman, C. and S. Sharma (2016). The power to choose: Gender balance of power and intra-household educational spending in India. WIDER Working Paper 2016/61, UNU-WIDER, Helsinki, Finland. ISBN: 9292561049.

OECD (2018). The Belt and Road Initiative in the global trade, investment and finance landscape. Technical report, OECD, Paris.

Oppenheimer, V. K. (1997). Women's employment and the gain to marriage: The specialization and trading model. *Annual review of sociology 23*(1), 431–453.

Parikh, J. K., A. Sharma, C. Singh, and S. Neelakantan (2016). *Providing clean cooking fuel in India: Challenges and solutions*. International Institute for Sustainable Development.

Peng, W., Z. Hisham, and J. Pan (2010). Household level fuel switching in rural Hubei. *Energy for sustainable development 14*(3), 238–244. ISBN: 0973-0826 Publisher: Elsevier.

Pitt, M. M. (1985). Equity, externalities and energy subsidies The case of kerosine in Indonesia. *Journal of Development Economics 17*(3), 201–217. ISBN: 0304-3878 Publisher: Elsevier.

Polinsky, A. M. and D. T. Ellwood (1979). An empirical reconciliation of micro and grouped estimates of the demand for housing. *The Review of Economics and Statistics*, 199–205.

Rao, M. N. and B. S. Reddy (2007). Variations in energy use by Indian households: an analysis of micro level data. *Energy 32*(2), 143–153. ISBN: 0360-5442 Publisher: Elsevier.

Reed, T. and A. Trubetskoy (2019). *Assessing the Value of Market Access from Belt and Road Projects*. The World Bank.

Rogers, W. and A. Winkler (2013, August). The relationship between the housing and labor market crises and doubling up: an MSA-level analysis, 2005–2011. *Monthly Labor Review*.

Rogers, W. H. and A. E. Winkler (2014). How Did the Housing and Labor Market Crises Affect Young Adults' Living Arrangements? IZA Discussion Paper 8568, IZA.

Ruiz-Mercado, I. and O. Masera (2015). Patterns of stove use in the context of fuel–device stacking: rationale and implications. *EcoHealth 12*(1), 42–56. ISBN: 1612-9202 Publisher: Springer.

Sachar, R., S. Hamid, T. K. Oommen, M. A. Basith, R. Basant, A. Majeed, and A. Shariff (2006). Social, economic and educational status of the Muslim community of India. Development Economics Working Papers 22136, East Asian Bureau of Economic Research.

Schaller, J. (2016, January). Booms, Busts, and Fertility: Testing the Becker Model Using Gender-Specific Labor Demand. *Journal of Human Resources 51*(1), 1–29.

Sehjpal, R., A. Ramji, A. Soni, and A. Kumar (2014). Going beyond incomes: Dimensions of cooking energy transitions in rural India. *Energy 68*, 470–477. ISBN: 0360-5442 Publisher: Elsevier.

Settersten Jr, R. A. and B. Ray (2010). What's going on with young people today? The long and twisting path to adulthood. *The future of children*, 19–41.

Sheiner, L. (1995). Housing prices and the savings of renters. *Journal of Urban Economics 38*(1), 94–125.

Shiller, R. (2008). Understanding Recent Trends in House Prices and Home Ownership. *Housing, Housing Finance and Monetary Policy* (Jackson Hole Conference Series, Federal Reserve Bank of Kansas City), 85–123.

Sinai, T. M. (2012). House price moments in boom-bust cycles. NBER Working Paper 18059, National Bureau of Economic Research.

Sironi, M. (2018, March). Economic Conditions of Young Adults Before and After the Great Recession. *Journal of Family and Economic Issues 39*(1), 103–116.

Sironi, M. and F. F. Furstenberg (2012, December). Trends in the Economic Independence of Young Adults in the United States: 1973-2007. *Population and Development Review 38*(4), 609–630.

Skinner, J. S. (1994). Housing and Saving in the United States. *NBER Chapters*, 191–214.

Smets, F. (2007). Commentary: Housing is the business cycle. *Housing, Housing Finance, and Monetary Policy*, 235–243.

Smith, K. R. (2000). National burden of disease in India from indoor air pollution. *Proceedings of the National Academy of Sciences 97*(24), 13286–13293. ISBN: 0027-8424 Publisher: National Acad Sciences.

Sun, C., N. Zhang, and Y. Liu (2017). One Belt One Road Initiatives and the Export Growth of China to the Related Countries. *Journal of International Trade* (2), 8.

Van der Kroon, B., R. Brouwer, and P. J. Van Beukering (2013). The energy ladder: Theoretical myth or empirical truth? Results from a meta-analysis. *Renewable and Sustainable Energy Reviews 20*, 504–513. ISBN: 1364-0321 Publisher: Elsevier.

Viswanathan, B. and K. K. Kumar (2005). Cooking fuel use patterns in India: 1983–2000. *Energy Policy 33*(8), 1021–1036. ISBN: 0301-4215 Publisher: Elsevier.

Wang, M. L., Q. Qiu, and C. H. Choi (2019, January). How will the Belt and Road initiative advance China's exports? *Asia Pacific Business Review 25*(1), 81–99.

World Health Organization (2011). *Health in the green economy: health co-benefits of climate change mitigation-housing sector*. World Health Organization.

Xu, Y. (2017, January). Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. *Political Analysis 25*(1), 57–76.

Zhang, Y., J.-H. Zhang, Q. Tian, Z.-H. Liu, and H.-L. Zhang (2018, May). Virtual water trade of agricultural products: A new perspective to explore the Belt and Road. *Science of The Total Environment 622-623*, 988–996.

Figures

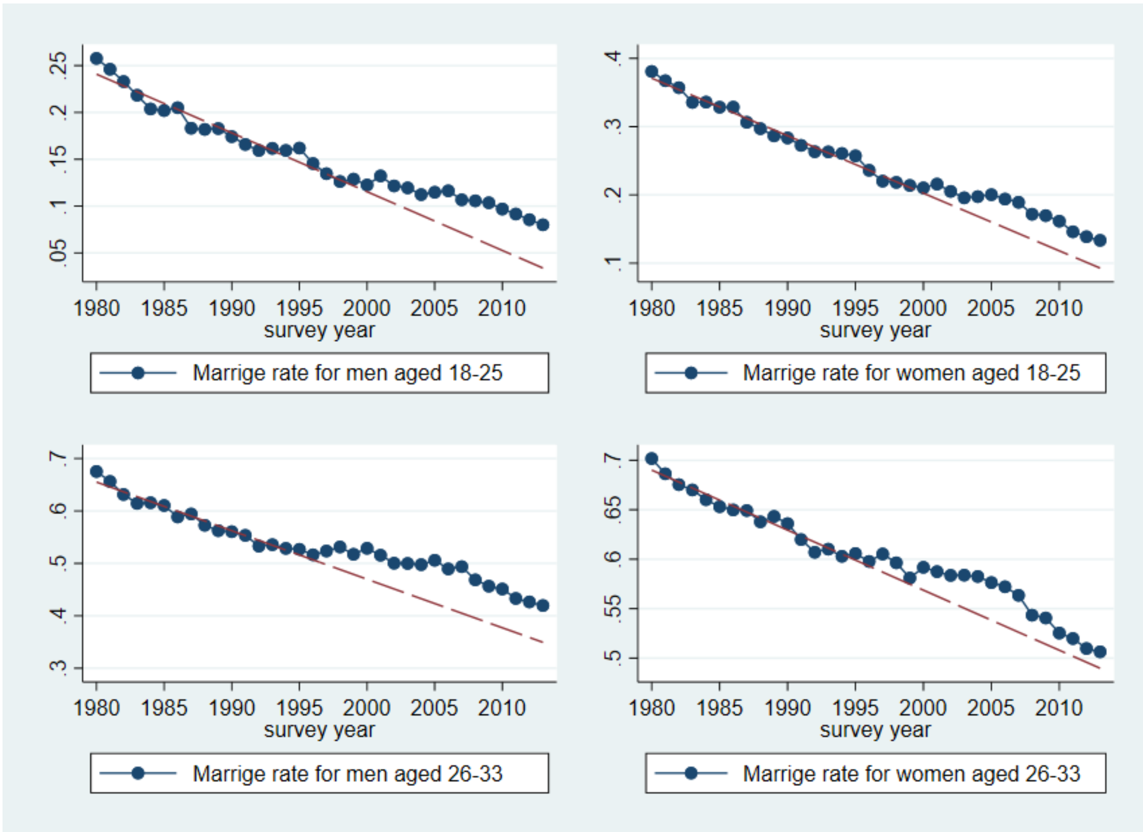


Figure 1: Share of Married People Age 18-25 and 26-33 by Gender, 1980-2013

Notes: This figure reports the married share of men and women age 18-25 and 26-33 during 1980-2013. The series are generated by the CPS survey using CPS person weight.The dashed line is the predicted marriage rate based on linear trend and it is fit to the 1980-1996 period

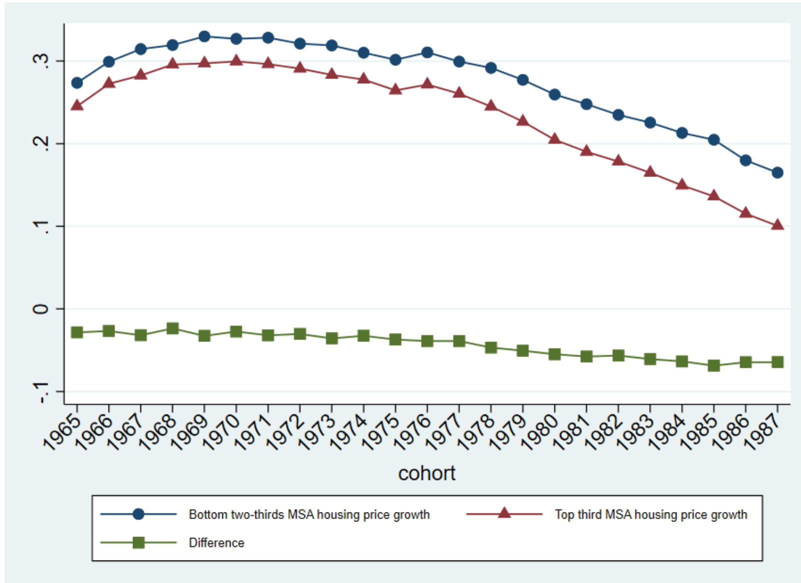


Figure 2: Share of Married People by Birth Cohort and MSA Housing Price Growth

Notes: This figure shows the marriage rate of people born between 1965 to 1987, which constructed by the 1990 and 2000 Censuses data and 2005-2013 ACS data. The sample is restricted to people age 25 to 54 in the survey year. The red-triangle line reports the marriage rate of people living in the MSAs belonging to the top tercile of housing price growth during 2000-2006. The blue-circle line is the marriage rate of people in the MSAs belonging to the relative bottom two terciles.

Figure 3: Fertility Rate by Birth Cohort and MSA Housing Price Growth

Notes: This figure shows the birth rate of people born between 1965 to 1987, which constructed by the 1990-2013 Vital Statistic data from NCHS. The sample is restricted to people age 25 to 48 in data year whose occurrence place is their residential place. The red-triangle line reports the fertility rate of people living in the MSAs at the top tercile of housing price growth during 2000-2006. The blue-circle line is the fertility rate of people in the MSAs with lower-tercile housing prices growth.

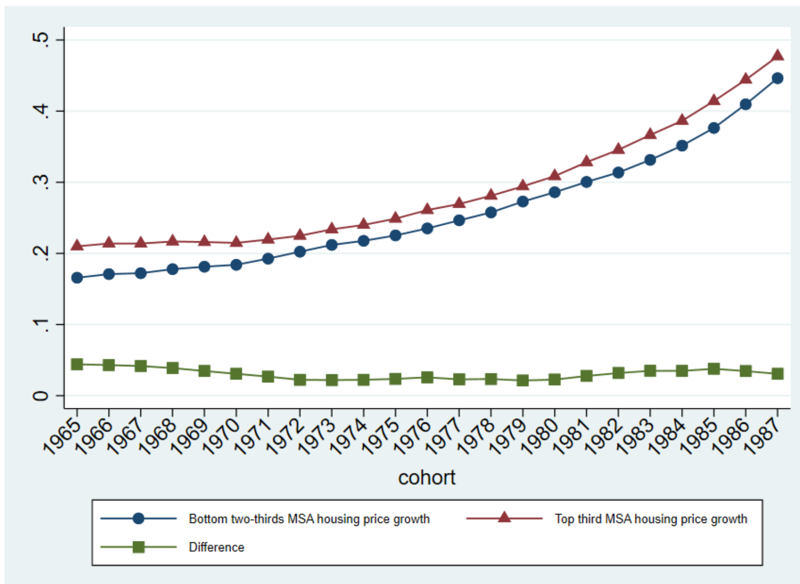


Figure 4: Nonmarital Birth Rate by Birth Cohort and MSA Housing Price Growth

Notes: This figure shows the nonmarital birth rate of people born between 1965 to 1987, which constructed by the 1990-2013 Vital Statistic data from NCHS. The sample is restricted to people age 25 to 48 in data year whose occurrence place is their residential place. The red-triangle line reports the nonmarital birth rate of people living in the MSAs at the top tercile of housing price growth during 2000-2006. The blue-circle line is the nonmarital birth rate of people in the MSAs with lower-tercile housing prices growth.

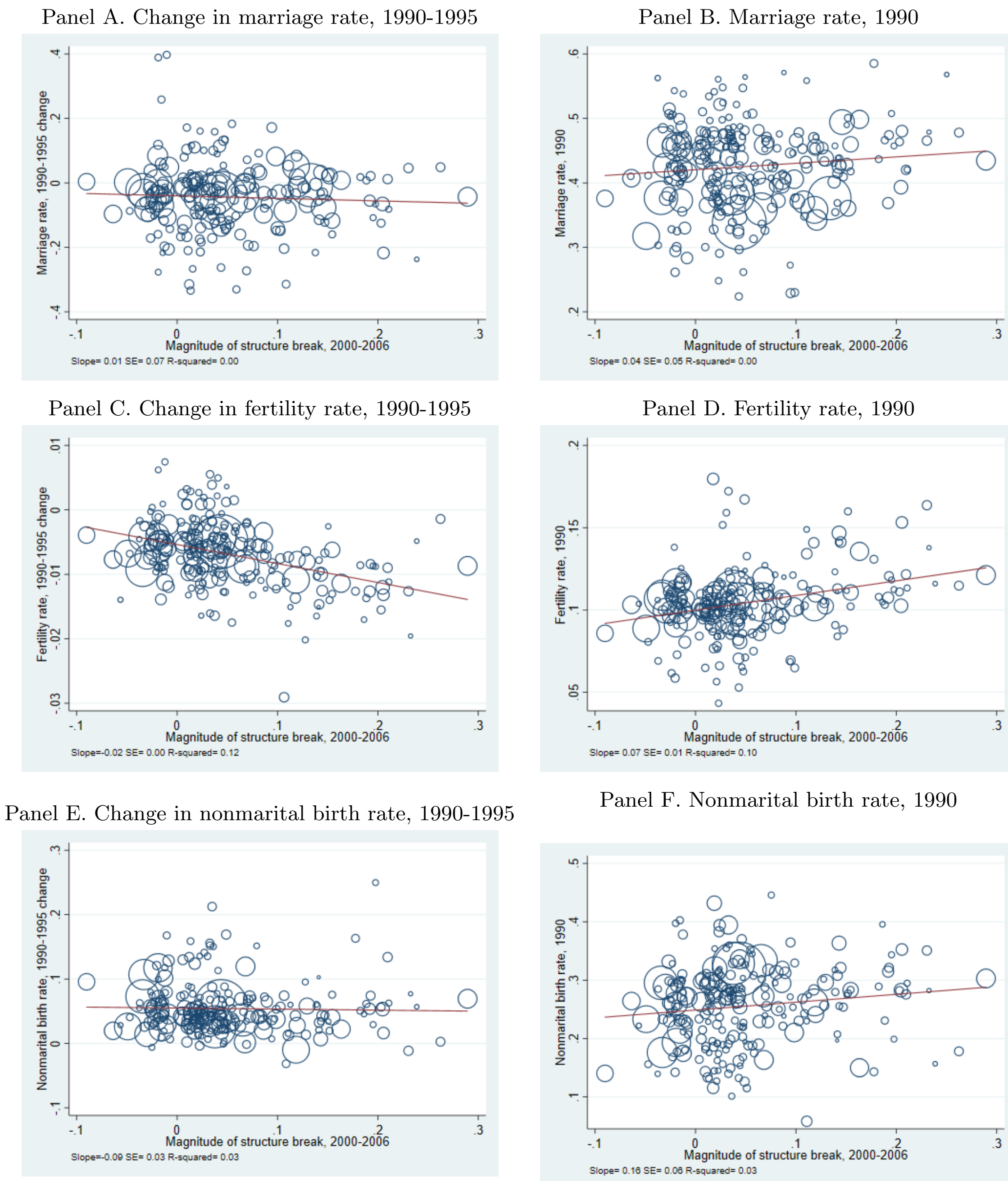


Figure 5: Correlation between Structure Break Instrument and Pre-period Variables

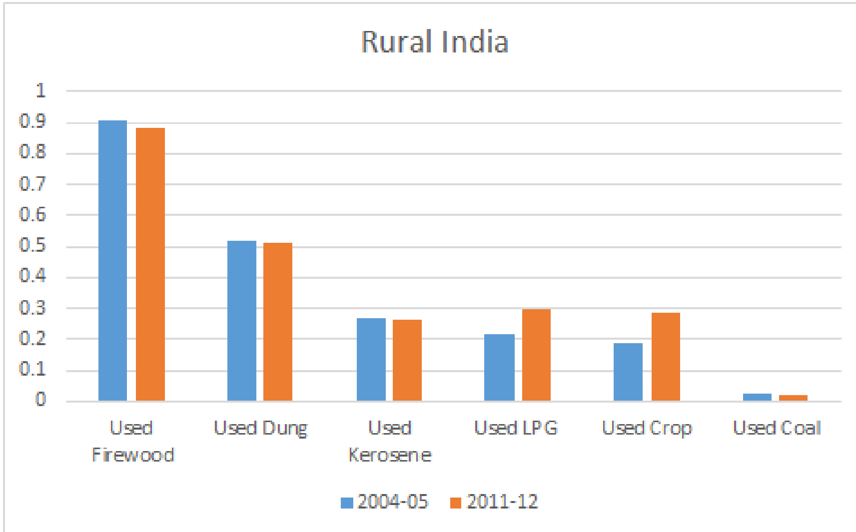


Figure 6: Types of Fuel Used by Rural Households by Survey Year

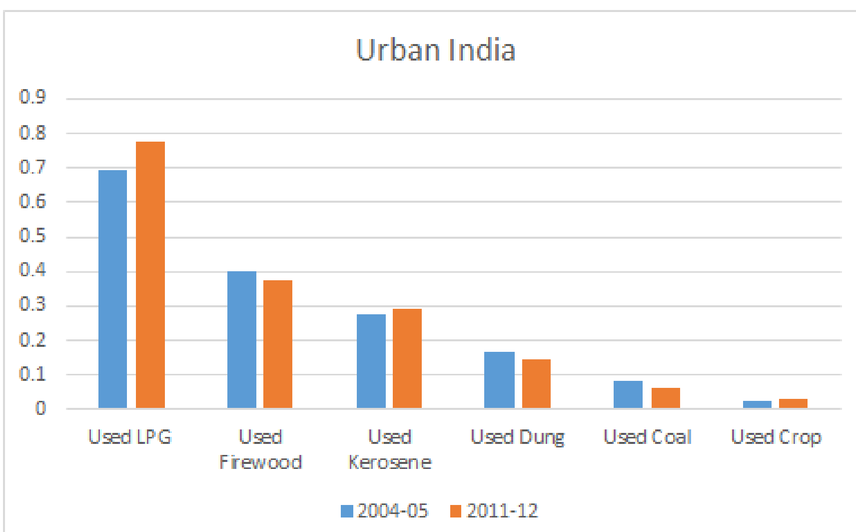


Figure 7: Types of Fuel Used by Urban Households by Survey Year

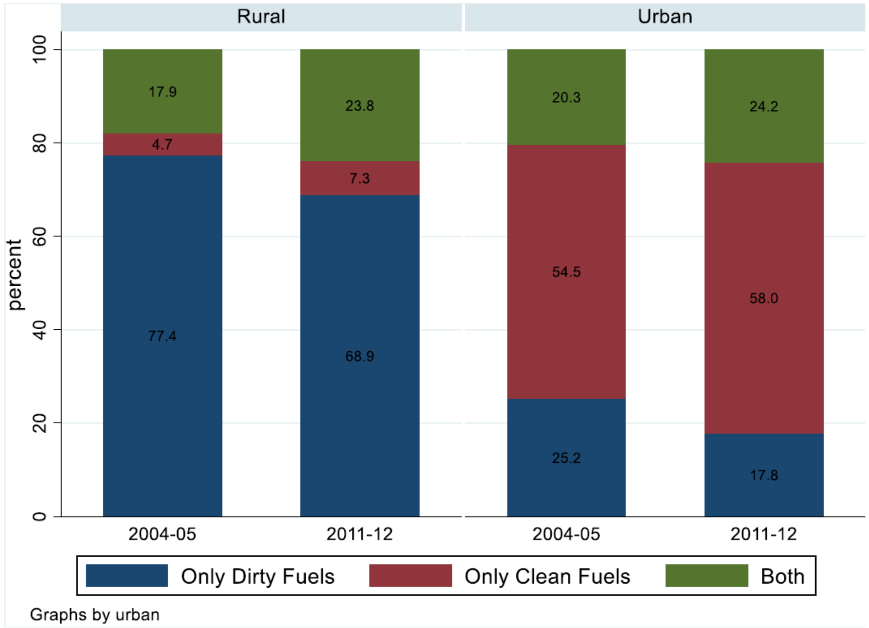


Figure 8: Fuel Stacking by Region and Survey Year

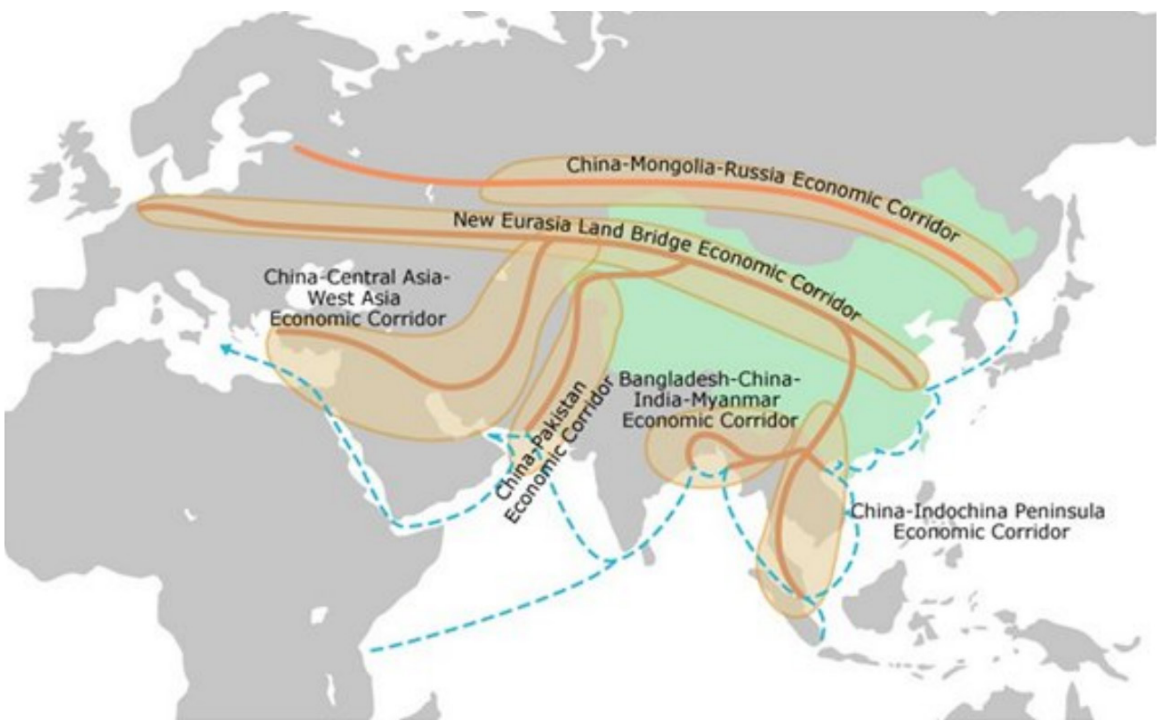


Figure 9: A Blueprint of the Belt and Road Initiative

Source: Hong Kong Trade Development Council (HKTDC) Research.

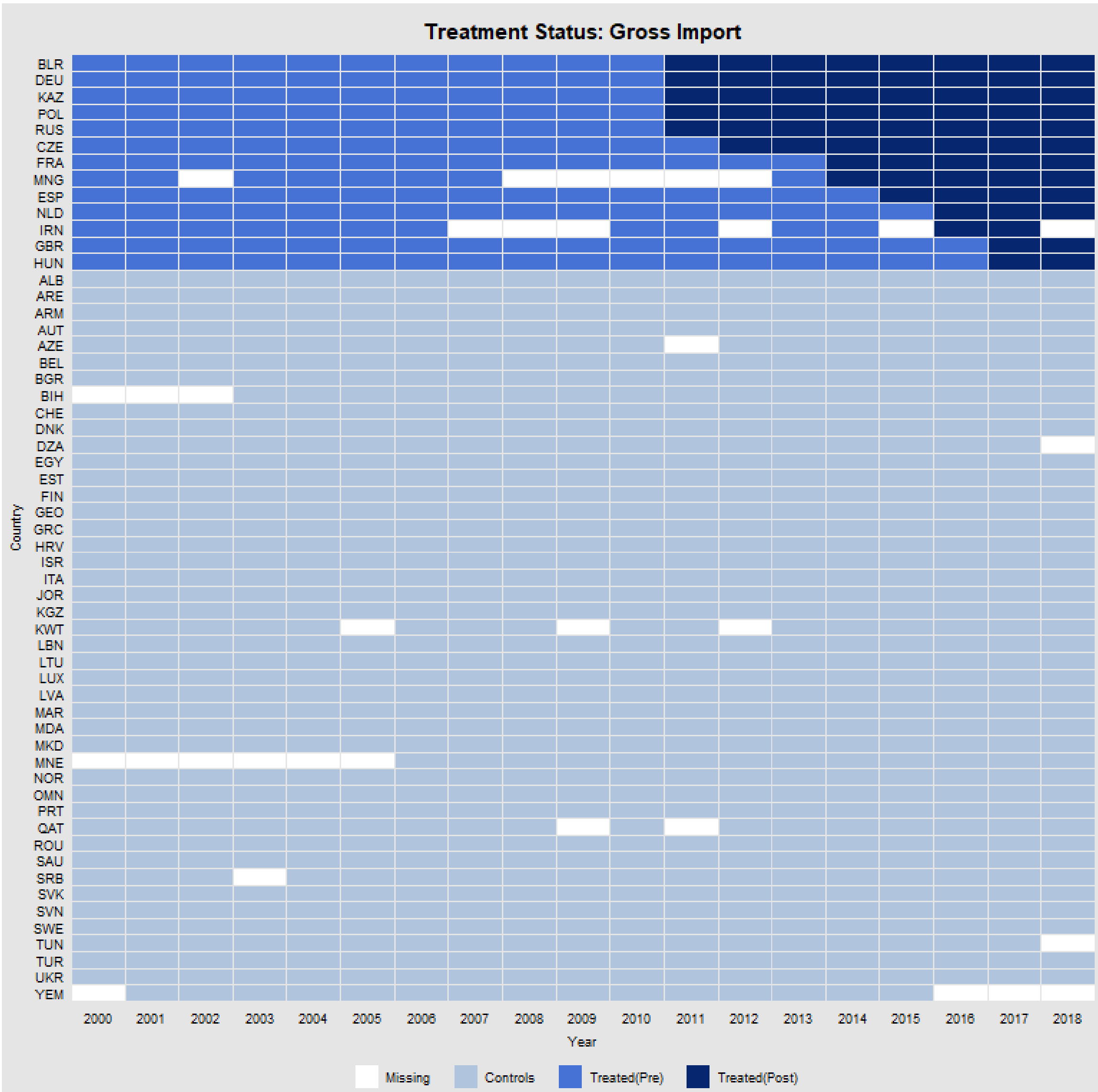
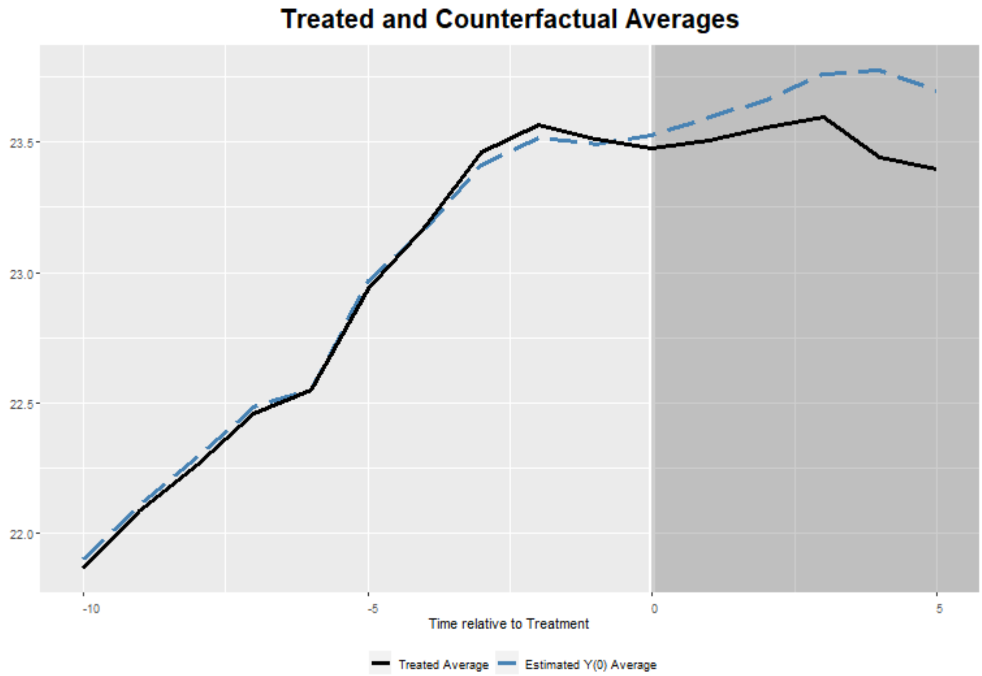
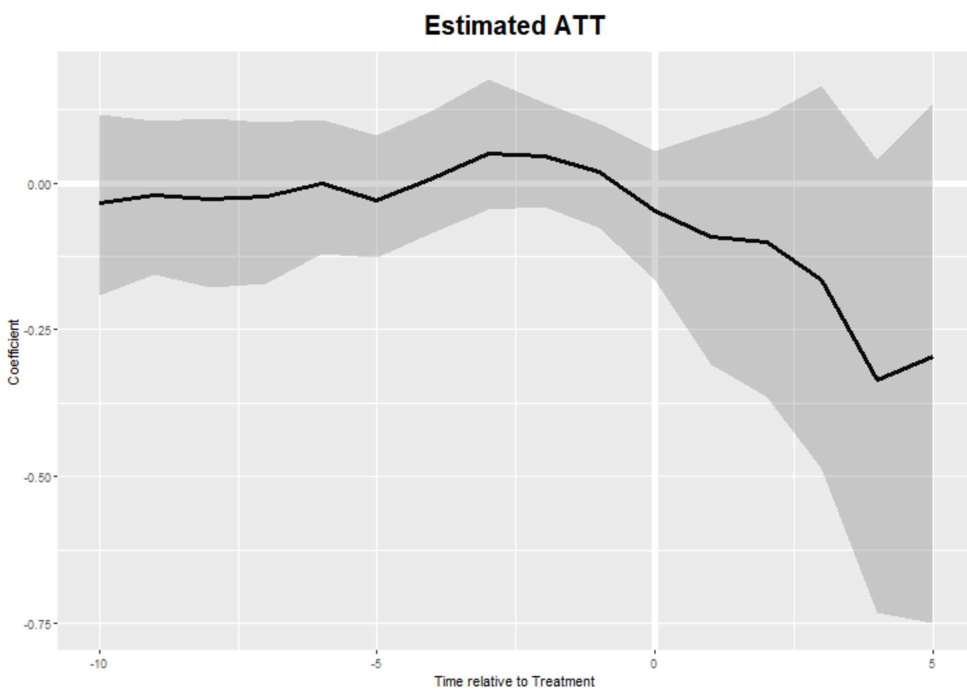


Figure 10: Data Structure of Gross Import



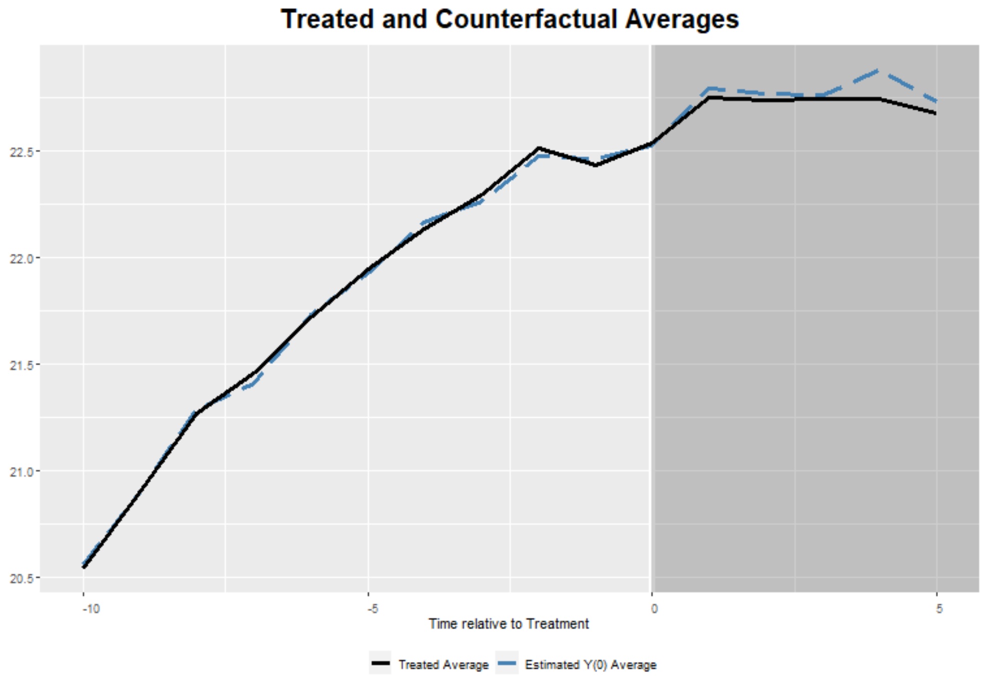
(a) Treated and Counterfactual Averages: Gross Import



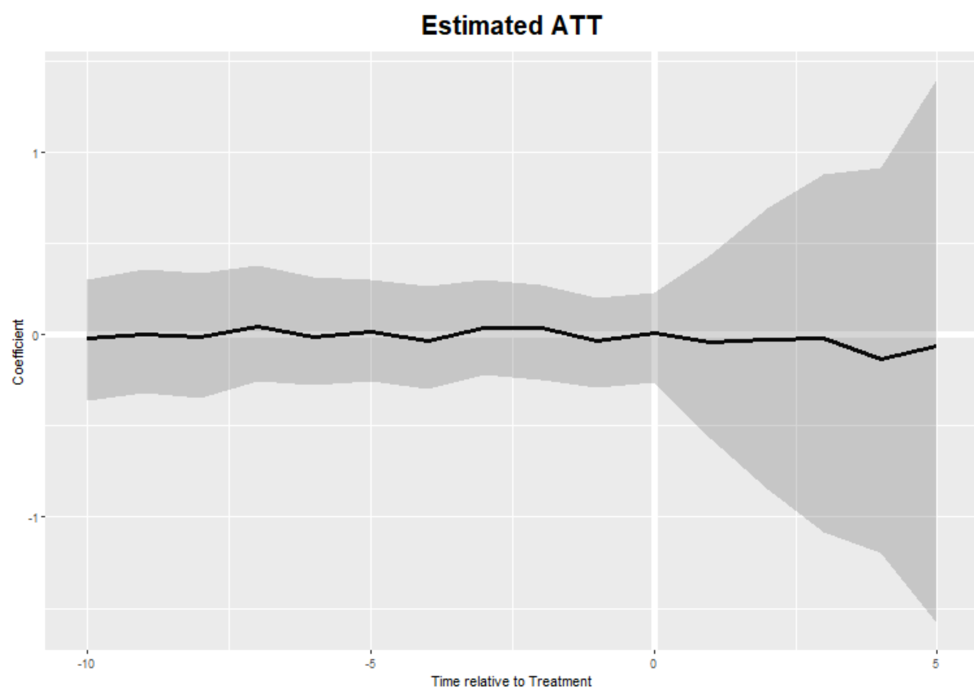
(b) Estimated ATT: Gross Import

Figure 11: The Effect of Railways on Gross Import by GSCM

Notes: The figure presents GSC estimates of the effect of New Silk Road Railways on gross import, which estimated by the baseline GSC model. Covariates, state and year fixed effects are included. In part (b), the confidence intervals are generated by block bootstrap of 1,000 times. The cross-validation scheme finds one unobserved factor.



(a) Treated and Counterfactual Averages: Gross Export



(b) Estimated ATT: Gross Export

Figure 12: The Effect of Railways on Gross Import by GSCM

Notes: The figure presents GSC estimates of the effect of New Silk Road Railways on gross export. Covariates, state and year fixed effects are included. In part (b), the confidence intervals are generated by block bootstrap of 1,000 times. The cross-validation scheme finds three unobserved factors.

Table 1: Descriptive Statistics of 2000-2006 changes by MSA

Variable	N	Mean	SD	Percentiles 10th	25th	50th	75th	90th	Source
Housing demand shock:	275	0.507	0.464	-0.089	0.158	0.494	0.822	1.085	
Change in housing prices	275	0.347	0.252	0.059	0.125	0.310	0.576	0.719	FHFA
Change in housing permits	275	0.160	0.302	-0.243	-0.028	0.154	0.328	0.524	Charles et al. (2018)
Structure Break	275	0.048	0.067	-0.03	-0.010	0.038	0.094	0.133	
Log of the age 18-55 population	275	13.732	1.391	11.831	12.673	13.697	14.808	15.752	ACS/Census
Female employment rate	275	0.694	0.052	0.638	0.661	0.689	0.727	0.755	ACS/Census
Employed with a college degree	275	0.218	0.066	0.133	0.175	0.214	0.265	0.309	ACS/Census
Foreign born share	275	0.008	0.004	0.004	0.005	0.007	0.009	0.013	ACS/Census
Change in employment rate:									
Age 18-25	275	-0.027	0.030	-0.059	-0.049	-0.033	-0.010	0.013	ACS/Census
Age 26-33	275	0.003	0.025	-0.026	-0.012	0.003	0.019	0.028	ACS/Census
Change in average wage									
Age 18-25	275	-0.011	0.056	-0.073	-0.038	-0.009	0.013	0.055	ACS/Census
Age 26-33	275	0.047	0.048	-0.006	0.017	0.036	0.081	0.100	ACS/Census
Change in marriage rate:									
Age 18-25	275	-0.040	0.021	-0.066	-0.048	-0.038	-0.028	-0.021	ACS/Census
Age 26-33	275	-0.065	0.027	-0.096	-0.076	-0.067	-0.051	-0.034	ACS/Census
Change in fertility rate:									
Age 18-25	235	-0.021	0.02	-0.036	-0.026	-0.018	-0.011	-0.007	Vital Statistics
Age 26-33	235	-0.022	0.02	-0.038	-0.03	-0.022	-0.013	0.004	Vital Statistics
Change in nonmarital birth:									
Age 18-25	235	0.098	0.04	0.057	0.077	0.095	0.126	0.142	Vital Statistics
Age 26-33	235	0.087	0.03	0.051	0.064	0.092	0.103	0.125	Vital Statistics

This is the summary statistic of the baseline sample at the MSA level. All of reported sample statistics are weighted by the 2000 population age 18-33 in the MSA.

Table 2: First-Stage Estimates

Dependent variable	Housing demand shock
Magnitude of structure break	3.833***
	(0.318)
Log of the age 18-55 population	0.006
	(0.019)
Female employment rate	-3.657***
	(0.410)
Employed with a college degree	1.525***
	(0.415)
Foreign born share	6.075
	(5.600)
Constant	2.401***
	(0.410)
Observations	275
R-squared	0.570

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 3: Housing Booms and Labor Market Outcomes

Dependent variable: 2000-2006 change in	Emp. Rate	Avg. wage	Emp. Rate	Avg. wage
	People Age 18-25		People Age 26-33	
Panel A. OLS estimates				
All people				
Housing demand change 2000-2006	0.022***	0.064***	0.015***	0.034***
	(0.006)	(0.014)	(0.005)	(0.010)
R2	0.331	0.255	0.326	0.243
Men				
Housing demand change 2000-2006	0.025***	0.081***	0.022**	0.042***
	(0.007)	(0.018)	(0.008)	(0.011)
R2	0.235	0.267	0.206	0.201
Women				
Housing demand change 2000-2006	0.033***	0.040***	0.004	0.018
	(0.008)	(0.010)	(0.008)	(0.011)
R2	0.104	0.067	0.041	0.144
Panel B. 2SLS estimates				
All people				
Housing demand change 2000-2006	0.030***	0.080***	0.019**	0.026
	(0.009)	(0.016)	(0.008)	(0.020)
Men				
Housing demand change 2000-2006	0.036***	0.107***	0.021*	0.033
	(0.012)	(0.022)	(0.011)	(0.022)
Women				
Housing demand change 2000-2006	0.041***	0.032**	0.019	0.007
	(0.015)	(0.014)	(0.030)	(0.018)
First stage F-statistic	48.940	48.940	48.940	48.940
Obs.	275	275	275	275
Baseline controls	Yes	Yes	Yes	Yes

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 4: Housing Booms and Marriage Outcomes: 2SLS Estimates, ACS/Census Data

Dependent variable: 2000-2006 change in share	Married	Cohabitating	Divorce	Never married	Married
Panel A. 2SLS estimates for people age 18-25					
Housing demand change 2000-2006	0.016*	0.001	0.001	-0.011	0.016*
	(0.008)	(0.004)	(0.001)	(0.011)	(0.008)
Average for people age 18-25 in 2000	0.147	0.090	0.014	0.710	0.147
Average for people age 18-25 in 2006	0.108	0.087	0.010	0.762	0.108
Panel B. 2SLS estimates for men age 18-25					
Housing demand change 2000-2006	0.018**	0.000	0.002*	-0.015	0.018**
	(0.008)	(0.004)	(0.001)	(0.010)	(0.008)
Average for men age 18-25 in 2000	0.113	0.078	0.010	0.767	0.113
Average for men age 18-25 in 2006	0.084	0.075	0.007	0.809	0.084
Panel C. 2SLS estimates for women age 18-25					
Housing demand change 2000-2006	0.013	0.001	0.000	-0.004	0.013
	(0.010)	(0.006)	(0.002)	(0.012)	(0.009)
Average for women age 18-25 in 2000	0.181	0.102	0.018	0.654	0.181
Average for women age 18-25 in 2006	0.131	0.100	0.014	0.714	0.131
Panel D. 2SLS estimates for people age 26-33					
Housing demand change 2000-2006	0.005	-0.009	0.001	0.001	0.005
	(0.006)	(0.006)	(0.004)	(0.003)	(0.007)
Average for people age 26-33 in 2000	0.518	0.097	0.066	0.265	0.518
Average for people age 26-33 in 2006	0.452	0.111	0.056	0.321	0.452
First stage F-statistic	48.940	48.940	48.940	48.940	48.940
Obs.	275	275	275	275	275
Baseline controls	Yes	Yes	Yes	Yes	Yes
Ownership rate	No	No	No	No	Yes

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 5: Housing Booms and Fertility Outcomes: 2SLS Estimates, Vital Statistics Data

Dependent variable: change in rate of	Fertility	Nonmarital birth	Fertility	Nonmarital birth	Fertility	Nonmarital birth	Fertility	Nonmarital birth
	2000-2006		2000-2007		2001-2007		2000-2006	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. 2SLS estimates for women age 18-25								
Housing demand change 2000-2006	0.025*	-0.053**	0.023*	-0.044	0.005	-0.017	0.025*	-0.053**
	(0.013)	(0.026)	(0.013)	(0.029)	(0.007)	(0.034)	(0.013)	(0.024)
Average for women age 18-25 in 2000/2001	0.104	0.523	0.104	0.523	0.086	0.548	0.104	0.523
Average for women age 18-25 in 2006/2007	0.084	0.613	0.084	0.629	0.084	0.629	0.084	0.613
Averge home ownership for people 18-25 in 2000							0.530	0.530
Panel B. 2SLS estimates for women age 26-33								
Housing demand change 2000-2006	0.028**	-0.028*	0.028**	-0.027	0.008*	-0.012	0.026**	-0.029*
	(0.012)	(0.016)	(0.012)	(0.020)	(0.004)	(0.017)	(0.012)	(0.016)
Average for women age 26-33 in 2000/2001	0.105	0.175	0.105	0.175	0.086	0.193	0.105	0.175
Average for women age 26-33 in 2006/2007	0.092	0.254	0.090	0.267	0.090	0.267	0.092	0.254
Averge home ownership for people 26-33 in 2000							0.612	0.612
First stage F-statistic	46.055	46.055	46.055	46.055	46.055	46.055	46.055	46.055
Obs.	235	235	235	235	235	235	235	235
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ownership rate	No	No	No	No	No	No	Yes	Yes

standard errors in parentheses.

*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 6: Placebo Tests of Housing Booms and Marriage Rate: 2SLS Estimates, CPS Data

Dependent variable: 1990-1996 change in		Marriage rate	
Panel A. 2SLS estimates for people age	18-25	26-33	18-33
Housing demand change 2000-2006	0.024	-0.030	-0.013
	(0.033)	(0.033)	(0.030)
Average level ar start of period	0.208	0.574	0.407
Panel B. 2SLS estimates for men age			
Housing demand change 2000-2006	0.024	-0.053	-0.028
	(0.031)	(0.045)	(0.036)
Average level ar start of period	0.159	0.543	0.367
Panel C. 2SLS estimates for women age			
Housing demand change 2000-2006	0.024	-0.000	0.006
	(0.040)	(0.036)	(0.032)
Average level ar start of period	0.257	0.61	0.448
First stage F-statistic	45.860	45.860	45.860
Obs.	192	192	192
Baseline controls	Yes	Yes	Yes

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 7: Placebo Tests of Housing Booms and Fertility Rate: 2SLS Estimates

Dependent variable: 1990-1996 change in	Fertility rate	Nonmarital birth rate
Panel A. 2SLS estimates for people age 18-25		
Housing demand change 2000-2006	-0.003	-0.068
	(0.005)	(0.049)
Average level ar start of period	0.106	0.417
Panel B. 2SLS estimates for people age 26-33		
Housing demand change 2000-2006	-0.008***	-0.013
	(0.002)	(0.020)
Average level ar start of period	0.103	0.145
Panel C. 2SLS estimates for people age 18-33		
Housing demand change 2000-2006	-0.006**	-0.041
	(0.003)	(0.035)
Average level ar start of period	0.104	0.271
First stage F-statistic	44.096	44.096
Obs.	233	233
Baseline controls	Yes	Yes

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 8: Housing Busts and Marriage Rate: 2SLS Estimates

Dependent variable: change in	Marriage rate				
	(1)	(2)	(3)	(4)	(5)
Panel A. 2SLS estimates	2006-2012	2006-2009	2009-2012	2008-2012	2000-2012
All people age 18-25					
Housing demand change 2000-2006	0.015***	0.008	0.007	0.009*	0.031***
	(0.005)	(0.005)	(0.006)	(0.005)	(0.011)
Men age 18-25					
Housing demand change 2000-2006	0.015***	0.008*	0.007	0.009	0.033***
	(0.005)	(0.004)	(0.006)	(0.005)	(0.011)
Women age 18-25					
Housing demand change 2000-2006	0.016**	0.007	0.009	0.010	0.029**
	(0.007)	(0.006)	(0.008)	(0.007)	(0.012)
Panel B. 2SLS estimates					
All people age 26-33					
Housing demand change 2000-2006	-0.021	-0.007	-0.013	-0.017	-0.016
	(0.013)	(0.010)	(0.010)	(0.011)	(0.011)
Men age 26-33					
Housing demand change 2000-2006	-0.015	-0.003	-0.012	-0.012	-0.013
	(0.019)	(0.011)	(0.015)	(0.017)	(0.016)
Women age 26-33					
Housing demand change 2000-2006	-0.026***	-0.011	-0.015*	-0.022***	-0.018**
	(0.009)	(0.011)	(0.009)	(0.008)	(0.009)
First stage F-statistic	48.940	48.940	48.940	48.940	48.940
Obs.	275	275	275	275	275
Baseline controls	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 9: Housing Busts and Fertility Rate: 2SLS Estimates

Dependent variable: change in	Birth rate		Nonmarital birth rate	
Panel A. 2SLS estimates for women age 18-25	2006-2012	2000-2012	2006-2012	2000-2012
Housing demand change 2000-2006	-0.007	0.018	0.012	-0.042
	(0.004)	(0.012)	(0.019)	(0.038)
Panel B. 2SLS estimates for women age 26-33				
Housing demand change 2000-2006	-0.005***	0.023*	0.017**	-0.011
	(0.002)	(0.012)	(0.008)	(0.021)
First stage F-statistic	46.055	46.055	46.055	46.055
Obs.	235	235	235	235
Baseline controls	Yes	Yes	Yes	Yes

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 10: Summary Statistics

	Rural				Urban			
	2004		2011		2004		2011	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Fuel use								
Firewood	0.906	0.292	0.884	0.320	0.398	0.489	0.381	0.486
Dung	0.509	0.500	0.536	0.499	0.168	0.374	0.147	0.354
Crop	0.184	0.387	0.293	0.455	0.026	0.158	0.030	0.170
Kerosene	0.255	0.436	0.267	0.442	0.283	0.450	0.299	0.458
LPG	0.210	0.408	0.298	0.458	0.684	0.465	0.781	0.414
Coal	0.026	0.160	0.017	0.130	0.081	0.274	0.061	0.238
Fuel choices								
Only Dirty Fuel	0.766	0.423	0.688	0.463	0.247	0.432	0.173	0.379
Only Clean Fuel	0.046	0.209	0.069	0.253	0.543	0.498	0.573	0.495
Mixed	0.188	0.391	0.243	0.429	0.210	0.407	0.254	0.435
Household characteristics								
Scheduled Caste/Tribes (SC/ST)	0.332	0.471	0.343	0.475	0.211	0.408	0.218	0.413
Muslim	0.092	0.290	0.091	0.288	0.149	0.356	0.162	0.368
Other Backward Class (OBC)	0.360	0.480	0.358	0.480	0.314	0.464	0.306	0.461
Household size	6.524	3.163	5.470	2.296	5.787	2.637	5.226	2.195
Dependency ratio	0.788	0.625	0.758	0.586	0.656	0.579	0.596	0.518
Main income source: salary	0.121	0.326	0.125	0.331	0.400	0.490	0.404	0.491
Main income source: non-agriculture wage	0.164	0.370	0.234	0.424	0.214	0.410	0.240	0.427
Main income source: trade	0.103	0.303	0.088	0.284	0.298	0.458	0.278	0.448
Log real per capita expenditure	6.345	0.628	6.613	0.621	6.774	0.636	7.091	0.658
Female head	0.040	0.195	0.057	0.232	0.056	0.230	0.080	0.272
Age of head	46.906	12.936	47.725	12.873	45.867	12.282	49.051	11.967
Head's years of education	4.358	4.375	4.944	4.562	7.492	4.867	7.891	4.860
Fuel prices								
Log firewood price	0.311	0.742	1.194	0.676	0.394	0.348	1.187	0.379
Log coal price	0.011	1.160	0.999	0.971	0.717	1.354	1.693	1.173
Log dung price	0.206	0.927	0.513	0.836				
Log kerosene price	2.635	0.350	3.460	0.240	2.823	0.216	3.379	0.266
Log LPG price	5.755	0.209	6.086	0.133	5.696	0.038	6.030	0.059
Village level variables								
Nearest town distance (km)	14.007	10.639	13.927	11.194				
Village has no paved road	0.327	0.469	0.136	0.343				
Percentage of households electrified	68.569	32.910	79.057	26.637				
Village main fuel is clean	0.060	0.238	0.154	0.361				
Village main fuel is biomass	0.169	0.375	0.846	0.361				
Women bargaining power								
Education gap	-2.814	4.392	-2.465	4.755	-2.025	4.329	-1.505	4.476
Violence index	0.099	1.022	0.055	0.995	-0.166	0.948	-0.126	0.997
Most say in decision making	-0.111	0.886	-0.108	0.896	0.092	0.997	-0.010	0.976
Financial independence	-0.078	0.947	-0.084	0.964	0.227	1.070	0.136	1.007
Permission needed index	0.089	0.936	0.079	0.924	-0.121	1.057	0.006	1.021
Sample size	27,679		26,779		10,961		11,861	

Table 11: Multinomial Logit with Random Effects for Fuel Choice, Rural and Urban Households

	Rural		Urban	
	(1)	(2)	(3)	(4)
	Clean only	***Mixed***	***Clean only***	***Mixed***
SC/ST	-1.714***	-1.482***	-2.332***	-1.217***
	(0.096)	(0.061)	(0.139)	(0.121)
Muslim	-0.485***	-0.372***	-1.863***	-0.949***
	(0.126)	(0.078)	(0.146)	(0.127)
Other Backward Class (OBC)	-0.903***	-0.829***	-1.447***	-0.888***
	(0.081)	(0.054)	(0.125)	(0.113)
Household size	0.057***	0.156***	0.072***	0.187***
	(0.013)	(0.007)	(0.017)	(0.015)
Deopendency ratio	-0.194***	-0.285***	-0.231***	-0.174***
	(0.057)	(0.034)	(0.065)	(0.056)
Main income source: salary	2.101***	1.225***	2.072***	0.886***
	(0.083)	(0.055)	(0.130)	(0.111)
Main income source: non-agriculture wage	0.431***	-0.017	0.282**	-0.047
	(0.099)	(0.054)	(0.126)	(0.102)
Main income source: trade	2.077***	0.944***	1.667***	0.372***
	(0.093)	(0.063)	(0.130)	(0.111)
Log real per capita consumption expenditure	1.645***	1.271***	2.125***	1.240***
	(0.055)	(0.036)	(0.079)	(0.072)
Female head	0.457***	0.504***	0.542***	0.250**
	(0.100)	(0.062)	(0.116)	(0.101)
Age of head	0.029***	0.020***	0.024***	0.011***
	(0.002)	(0.002)	(0.003)	(0.003)
Head's years of education	0.244***	0.176***	0.309***	0.161***
	(0.008)	(0.005)	(0.011)	(0.010)
Log firewood price	0.191***	0.056**	1.073***	-0.150
	(0.041)	(0.025)	(0.117)	(0.102)
Log coal price	-0.070**	0.107***	-0.075**	0.269***
	(0.028)	(0.017)	(0.029)	(0.027)
Log dung price	-0.126***	-0.305***		
	(0.030)	(0.020)		
Log kerosene price	0.585***	0.518***	1.546***	0.411***
	(0.100)	(0.065)	(0.175)	(0.151)
Log LPG price	0.206	-0.262**	-2.359***	-1.537**
	(0.261)	(0.111)	(0.767)	(0.686)
Nearest town distance (km)	-0.007**	-0.010***		
	(0.003)	(0.002)		
Village has no paved road	-0.727***	-0.423***		
	(0.092)	(0.050)		
Percentage of households electrified	0.019***	0.017***		
	(0.002)	(0.001)		
Village main fuel is clean	1.317***	0.864***		
	(0.100)	(0.069)		
Village main fuel is biomass	-0.425***	-0.134**		
	(0.093)	(0.058)		
Constant	-22.040***	-12.881***	-6.935	-2.101
	(1.583)	(0.719)	(4.432)	(3.963)

Biomass fuel only is base category. Standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 11: Multinomial Logit with Random Effects for Fuel Choice, Rural and Urban Households (Cont.)

	Rural		Urban	
	(1)	(2)	(3)	(4)
	Clean only	***Mixed***	***Clean only***	***Mixed***
Heterogeneity covariance				
var(a1)	5.077		6.812	
	(0.333)		(0.474)	
var(a2)	3.066		3.276	
	(0.155)		(0.277)	
Covariance(a1,a2)	3.227***		3.309***	
	(0.196)		(0.317)	
Region controls	Yes		Yes	
Year controls	Yes		Yes	
Log likelihood	-25639.121		-15117.697	
Observations	52,409		21,911	

Biomass fuel only is base category. Standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 12: Average Marginal Effects of Multinomial Logit with Random Effects

	Rural			Urban		
	(1) Dirty Only	(2) Clean Only	(3) Mixed	(4) Dirty Only	(5) Clean Only	(6) Mixed
SC/ST	0.132***	-0.029***	-0.104***	0.129***	-0.156***	0.027***
	(0.005)	(0.003)	(0.005)	(0.008)	(0.009)	(0.009)
Muslim	0.034***	-0.009**	-0.025***	0.102***	-0.126***	0.025**
	(0.007)	(0.004)	(0.006)	(0.009)	(0.010)	(0.010)
Other Backward Class (OBC)	0.073***	-0.014***	-0.059***	0.086***	-0.087***	0.002
	(0.005)	(0.002)	(0.004)	(0.008)	(0.008)	(0.008)
Household size	-0.012***	-0.001***	0.013***	-0.010***	-0.006***	0.016***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Deopendency ratio	0.024***	-0.001	-0.023***	0.015***	-0.012**	-0.003
	(0.003)	(0.002)	(0.003)	(0.004)	(0.005)	(0.005)
Main income source: salary	-0.119***	0.048***	0.071***	-0.106***	0.153***	-0.047***
	(0.005)	(0.002)	(0.004)	(0.008)	(0.010)	(0.010)
Main income source: non-agriculture wage	-0.005	0.015***	-0.011**	-0.007	0.033***	-0.026**
	(0.005)	(0.003)	(0.005)	(0.008)	(0.011)	(0.010)
Main income source: trade	-0.098***	0.053***	0.045***	-0.070***	0.147***	-0.077***
	(0.005)	(0.003)	(0.005)	(0.008)	(0.011)	(0.010)
Log real per capita consumption expenditure	-0.116***	0.031***	0.085***	-0.123***	0.133***	-0.010*
	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)	(0.005)
Female head	-0.043***	0.005*	0.038***	-0.028***	0.039***	-0.010
	(0.005)	(0.003)	(0.005)	(0.007)	(0.009)	(0.009)
Age of head	-0.002***	0.001***	0.001***	-0.001***	0.002***	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Head's years of education	-0.016***	0.005***	0.011***	-0.017***	0.021***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Log firewood price	-0.007***	0.005***	0.001	-0.028***	0.123***	-0.094***
	(0.002)	(0.001)	(0.002)	(0.007)	(0.009)	(0.008)
Log coal price	-0.007***	-0.005***	0.011***	-0.009***	-0.027***	0.036***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Log dung price	0.024***	0.002**	-0.026***			
	(0.002)	(0.001)	(0.002)			
Log kerosene price	-0.046***	0.010***	0.037***	-0.068***	0.132***	-0.064***
	(0.005)	(0.003)	(0.005)	(0.011)	(0.013)	(0.013)
Log LPG price	0.016*	0.013	-0.029***	0.143***	-0.136**	-0.007
	(0.010)	(0.009)	(0.010)	(0.050)	(0.058)	(0.057)
Nearest town distance (km)	0.001***	-0.000	-0.001***			
	(0.000)	(0.000)	(0.000)			
Village has no paved road	0.041***	-0.017***	-0.025***			
	(0.004)	(0.003)	(0.004)			
Percentage of households electrified	-0.002***	0.000***	0.001***			
	(0.000)	(0.000)	(0.000)			
Village main fuel is clean	-0.082***	0.028***	0.054***			
	(0.006)	(0.003)	(0.006)			
Village main fuel is biomass	0.016***	-0.012***	-0.004			
	(0.005)	(0.003)	(0.005)			
Regional controls		Yes			Yes	
Year controls		Yes			Yes	
Observations		52,409			21,911	

Standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 13: Average Marginal Effects of the Multinomial Random Effects with Women Bargaining Power

	Rural			Urban		
	(1) Dirty Only	(2) Clean Only	(3) Mixed	(4) Dirty Only	(5) Clean Only	(6) Mixed
SC/ST	0.114***	-0.021***	-0.093***	0.103***	-0.136***	0.033***
	(0.006)	(0.003)	(0.006)	(0.010)	(0.010)	(0.010)
Muslim	0.005	0.001	-0.007	0.080***	-0.107***	0.027**
	(0.008)	(0.004)	(0.008)	(0.010)	(0.012)	(0.012)
Other Backward Class (OBC)	0.054***	-0.011***	-0.043***	0.070***	-0.074***	0.004
	(0.005)	(0.003)	(0.005)	(0.009)	(0.010)	(0.010)
Household size	-0.009***	-0.001**	0.010***	-0.006***	-0.009***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Deopendency ratio	0.027***	0.000	-0.027***	0.015***	-0.006	-0.009
	(0.004)	(0.002)	(0.004)	(0.005)	(0.007)	(0.007)
Main income source: salary	-0.117***	0.047***	0.070***	-0.115***	0.164***	-0.049***
	(0.006)	(0.003)	(0.005)	(0.010)	(0.013)	(0.012)
Main income source: non-agriculture wage	-0.006	0.011***	-0.005	-0.008	0.049***	-0.041***
	(0.005)	(0.004)	(0.006)	(0.009)	(0.013)	(0.012)
Main income source: trade	-0.092***	0.052***	0.039***	-0.073***	0.150***	-0.077***
	(0.006)	(0.003)	(0.006)	(0.010)	(0.013)	(0.012)
Log real per capita consumption expenditure	-0.118***	0.031***	0.086***	-0.117***	0.130***	-0.013**
	(0.004)	(0.002)	(0.004)	(0.006)	(0.006)	(0.007)
Female head	-0.053***	0.010**	0.043***	-0.054***	0.065***	-0.011
	(0.009)	(0.005)	(0.009)	(0.011)	(0.013)	(0.013)
Age of head	-0.003***	0.001***	0.002***	-0.002***	0.002***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Head's years of education	-0.018***	0.005***	0.012***	-0.017***	0.021***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Log firewood price	-0.006**	0.005***	0.001	-0.032***	0.129***	-0.096***
	(0.003)	(0.002)	(0.003)	(0.009)	(0.010)	(0.010)
Log coal price	-0.008***	-0.004***	0.011***	-0.011***	-0.028***	0.039***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Log dung price	0.028***	0.002**	-0.030***			
	(0.002)	(0.001)	(0.002)			
Log kerosene price	-0.051***	0.011***	0.040***	-0.064***	0.144***	-0.081***
	(0.007)	(0.004)	(0.007)	(0.013)	(0.015)	(0.015)
Log LPG price	0.020*	0.003	-0.023*	0.243***	-0.238***	-0.005
	(0.011)	(0.009)	(0.012)	(0.058)	(0.067)	(0.068)
Nearest town distance (km)	0.001***	-0.000	-0.001***			
	(0.000)	(0.000)	(0.000)			
Village has no paved road	0.037***	-0.018***	-0.020***			
	(0.005)	(0.004)	(0.005)			
Percentage of households electrified	-0.002***	0.000***	0.001***			
	(0.000)	(0.000)	(0.000)			
Village main fuel is clean	-0.072***	0.027***	0.045***			
	(0.007)	(0.004)	(0.007)			
Village main fuel is biomass	0.031***	-0.012***	-0.019***			
	(0.006)	(0.004)	(0.006)			
Education gap	-0.007***	0.002***	0.005***	-0.006***	0.008***	-0.002**
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Financial independence	-0.018***	0.004***	0.014***	-0.024***	0.034***	-0.011***
	(0.002)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)
Most say in decition making	-0.002	0.001	0.001	0.006**	0.001	-0.008**
	(0.002)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)
Violence	0.009***	0.004***	-0.013***	-0.001	0.010***	-0.010***
	(0.002)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)
Permission Index	-0.001	0.001	-0.000	-0.004	0.006*	-0.002
	(0.002)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)
Regional controls		Yes			Yes	
Year controls		Yes			Yes	
Observations		36,074			15,455	

Standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 14: Summary Statistics, 2000-2018

	N	Mean	SD	Min	Max
Treatment group					
Log of Gross export	236	21.840	1.721	17.520	25.430
Log of Gross import	235	22.880	1.736	17.540	25.570
Log of exchange rate	236	0.287	3.065	-4.548	8.500
Manufacturing/GDP	234	0.152	0.051	0.054	0.286
Agriculture/GDP	236	0.043	0.042	0.006	0.274
Labor force participation rate	236	0.596	0.071	0.371	0.821
Imp/GDP	236	0.441	0.188	0.176	0.811
Exp/GDP	236	0.463	0.191	0.193	0.880
Log of total population	236	17.140	1.060	14.690	18.800
Tariff rate	232	0.037	0.039	0.013	0.227
Log of GDP at 2010 USD	236	26.820	1.681	22.070	29.000
Control group					
Log of Gross export	814	19.050	2.626	6.922	24.130
Log of Gross import	814	20.910	1.714	13.990	24.570
Log of exchange rate	817	-0.557	1.978	-3.335	5.503
Manufacturing/GDP	804	0.154	0.131	0.037	1.615
Agriculture/GDP	805	0.058	0.057	0.001	0.345
Labor force participation rate	817	0.565	0.108	0.148	0.918
Imp/GDP	801	0.506	0.229	0.143	1.872
Exp/GDP	801	0.481	0.269	0.099	2.212
Log of total population	817	15.720	1.146	12.990	18.400
Tariff rate	802	0.040	0.040	0.003	0.271
Log of GDP at 2010 USD	817	25.080	1.503	21.890	28.440

Table 15: Regression Estimates of Railways on Gross Import

	FE			GSC	
Dep: log of gross import	(1)	(2)	(3)	(4)	(5)
Railway connection	0.041	-0.029	0.051	-0.134	-0.199*
	(0.149)	(0.095)	(0.046)	(0.118)	(0.115)
Log of GDP at 2010 USD		1.690***	1.401***		1.548***
		(0.299)	(0.350)		(0.345)
Log of total population		-0.603**	0.325		-0.385
		(0.297)	(0.502)		(0.354)
Labor force participation rate		0.023	-0.604		-0.047
		(0.630)	(0.502)		(0.589)
Exp/GDP		-1.245***	-1.243***		-1.139***
		(0.362)	(0.302)		(0.366)
Imp/GDP		1.384***	1.166***		1.209***
		(0.336)	(0.279)		(0.332)
Manufacturing/GDP		-0.844***	-0.597		-0.732
		(0.160)	(0.450)		(0.819)
Agriculture/GDP		-3.479*	-4.000**		-4.816**
		(1.817)	(1.757)		(2.193)
Log of exchange rate		0.083	-0.340***		0.046
		(0.059)	(0.088)		(0.117)
Tariff rate		-0.689	1.056*		-0.131
		(1.091)	(0.598)		(1.089)
Constant	19.374***	-13.030**	-17.868**		
	(0.100)	(5.742)	(7.882)		
Observations	1,049	1,002	1,002	1,053	1,053
R-squared	0.960	0.922	0.991	N/A	N/A
Country FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Time trends	No	No	Linear	N/A	N/A
Unobserved factors	N/A	N/A	N/A	3	1

Standard errors are robust and clustered at country level. Standard errors in column (4) and (5) are bootstraps (blocked at the country level) of 1,000 times.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 16: Regression Estimates of Railways on Gross Export

	FE			GSC	
Dep: log of gross export	(1)	(2)	(3)	(4)	(5)
Railway connection	-0.300	-0.446***	0.071	0.166	-0.107
	(0.183)	(0.135)	(0.114)	(0.448)	(0.417)
Log of GDP at 2010 USD		2.232***	0.826		2.509***
		(0.547)	(0.499)		(0.763)
Log of total population		-1.847**	-1.242		-2.188***
		(0.894)	(0.861)		(1.114)
Labor force participation rate		-1.275	-1.751		0.598
		(1.364)	(2.286)		(1.444)
Exp/GDP		2.444***	1.838***		1.749**
		(0.830)	(0.605)		(0.825)
Imp/GDP		-1.914*	-0.535		-1.068*
		(1.005)	(0.614)		(1.020)
Manufacturing/GDP		0.660	-2.419**		0.822
		(0.464)	(0.911)		(2.575)
Agriculture/GDP		4.850	5.511		2.113
		(3.072)	(5.540)		(4.347)
Log of exchange rate		-0.102	-0.123		0.008
		(0.209)	(0.191)		(0.388)
Tariff rate		-1.093	-0.926		0.021
		(2.711)	(1.989)		(3.313)
Constant	17.671***	-9.090	10.639		
	(0.199)	(18.121)	(13.021)		
Observations	1,050	1,004	1,004	1,053	1,053
R-squared	0.918	0.687	0.962	N/A	N/A
Country FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Time trends	No	No	Linear	N/A	N/A
Unobserved factors	N/A	N/A	N/A	4	3

Standard errors are robust and clustered at country level. Standard errors in column (4) and (5) are bootstraps (blocked at the country level) of 1,000 times.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

APPENDIX

Data Appendix

Matching FHFA and MSA geographic codes

There are 283 MSAs in my marriage data organized from a combined data set of 2000 census, 2005-2007, and 2009-2011 ACS. 259 of them can directly match to MSAs in FHFA HPI data. 8 of the remaining 24 MSAs are matched to the same FHFA MSAs. The last 16 MSAs are mappedto the nearest FHFA MSAs. The following table list 24 MSAs and its corresponding FHFA MSAs:

Census MSAs	FHFA MSAs
120 Brockton, Ma	14454 Boston, MA
380 Kenosha, WI	16974 Chicago-Naperville-Arlington Heights, IL
320 Hamilton-Middleton, OH	17140 Cincinnati, OH-KY-IN
275 Fort Walton Beach, FL	18880 Crestview-Fort Walton Beach-Destin, FL
202 Daytona Beach, FL	19660 Deltona-Daytona Beach-Ormond Beach, FL
361 Jamestown-Dunkirk, NY	21500 Erie, PA
87 Benton Harbor, MI	24340 Grand Rapids-Wyoming, MI
92 Biloxi-Gulfport, MS	24860 Greenville-Anderson-Mauldin, SC
292 Galveston-Texas City, TX	26420 Houston-The Woodlands-Sugar Land, TX
260 Fitchburg-Leominster, MA	31700 Manchester-Nashua, NH
519 Monmouth-Ocean, NJ	35084 Newark, NJ-PA
548 New Haven-Meriden, CT	35300 New Haven-Milford, CT
566 Newburgh-Middletown, NY	35614 New York-Jersey City-White Plains, NY-NJ
751 Sarasota, FL	35840 North Port-Sarasota-Bradenton, FL
873 Ventura-Oxnard-Simi Valley, CA	37100 Oxnard-Thousand Oaks-Ventura, CA
490 Melbourne-Titusville-Cocoa-Palm Bay, FL	37340Palm Bay-Melbourne-Titusville, FL
271 Fort Pierce, FL	38940 Port St. Lucie, FL
540 New Bedford, MA	39300 Providence-Warwick, RI-MA
927 Yolo, CA	40900 Sacramento–Roseville–Arden-Arcade, CA
761 Sharon, PA	49660 Youngstown-Warren-Boardman, OH-PA
535 Nashua, NH	31700 Manchester-Nashua, NH
193 Danbury, CT	35300 New Haven-Milford, CT
804 Stamford, CT	35300 New Haven-Milford, CT
888 Waterbury, CT	35300 New Haven-Milford, CT

Appendix Graphs

Figure A1: Correlation between Changes in Housing Permits and Home Prices, 2000-2006

Notes: This figure is the replication of Figure 6 in Charles, Hurst, and Notowidigdo(2018). The red line is the regression weighted by 18-55 adult population in 2000, and the sample includes 275 MSAs. Yuba city has extraordinary growth in housing permit, so it is excluded from the figure.

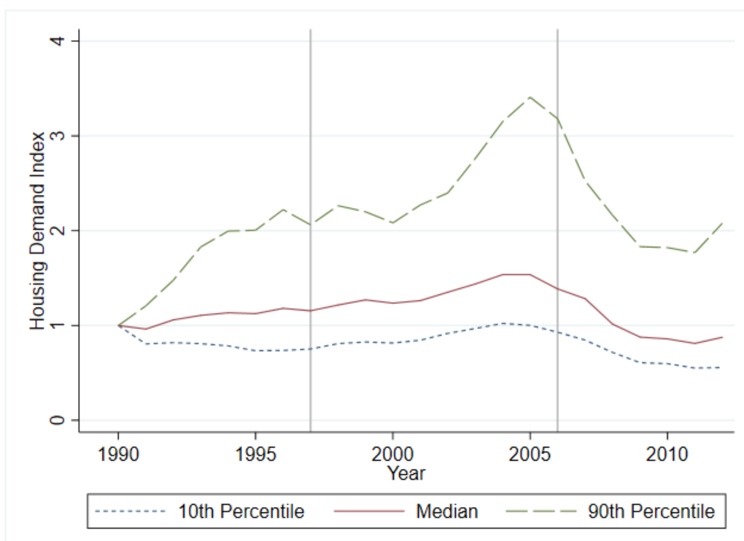


Figure A2: Time Variation in Housing Demand across MSAs

Notes: This figure is the replication of Figure 7 in Charles, Hurst, and Notowidigdo(2018). The lines are time series of the housing demand index (HDI) at the tenth percentile, median, and ninetieth percentile. The HDI is the average sum of prices and permits, which is normalized to 1990 values by each percentile.

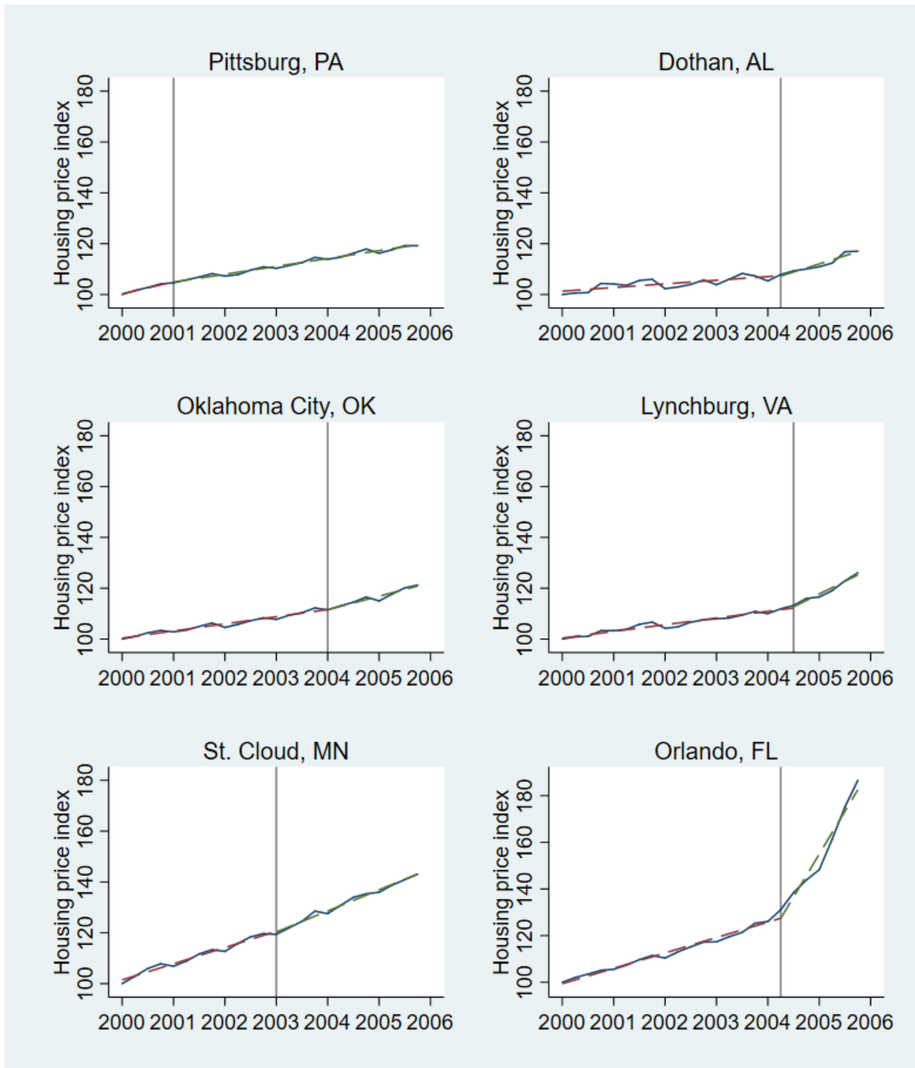


Figure A3: Variation in Structure Break across MSAs

Notes: This graphs include quarterly normalized HPI for six MSAs between 2000:I and 2005:I. The solid line is HPI, and the dashed line is fitted line before and after the structure break.The grey vertical line shows the time may have structure break. The first column include three MSAs with a smaller estimated structure break, while the second column includes the MSAs with a larger structure break.

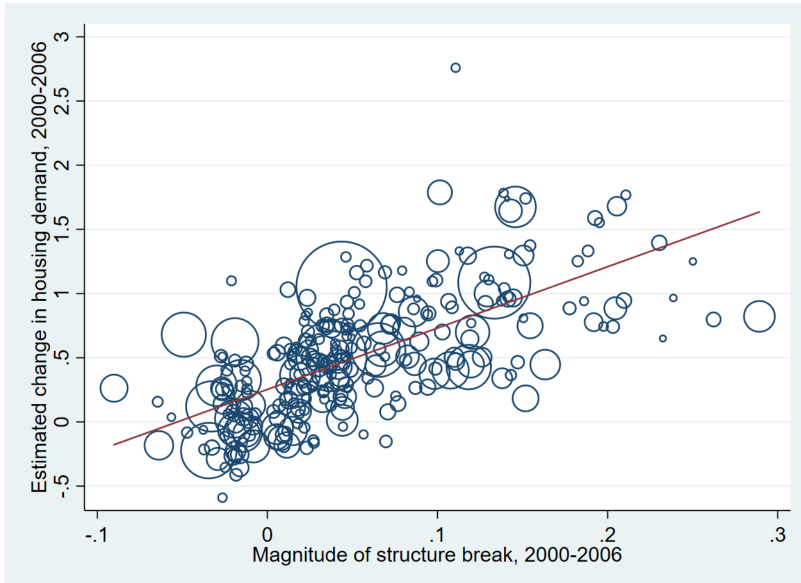


Figure A4: Relationship between Instrument and Change in Housing Demand

Notes: This figure reports the correlation between the magnitude of structural break and the estimated housing demand change during 2000-2006.

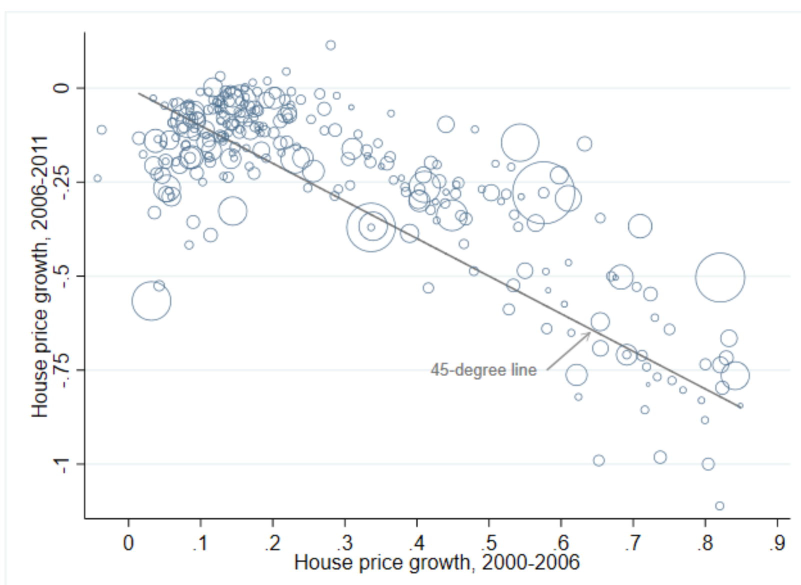
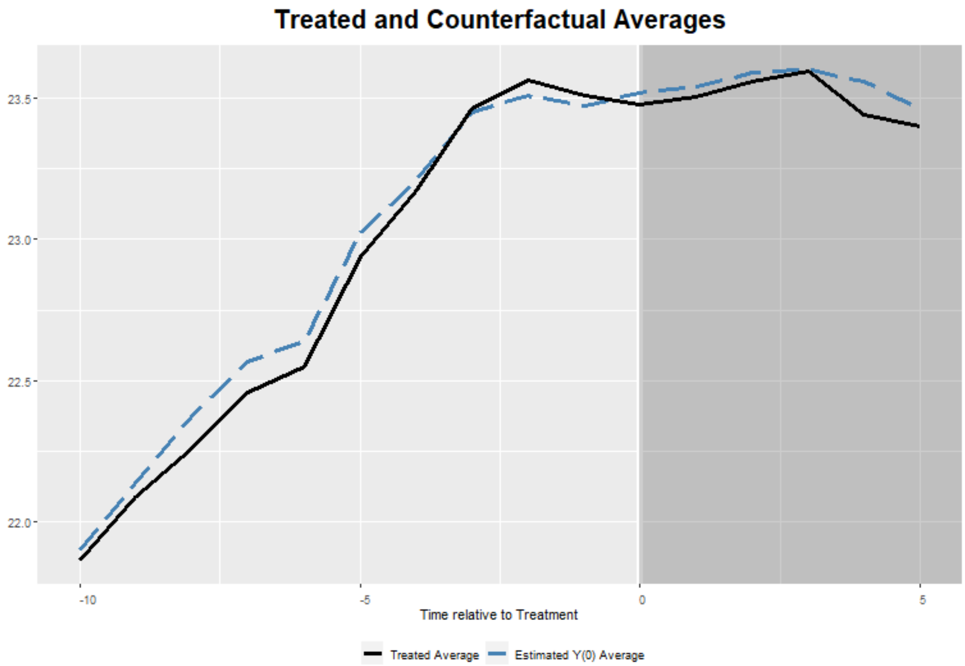
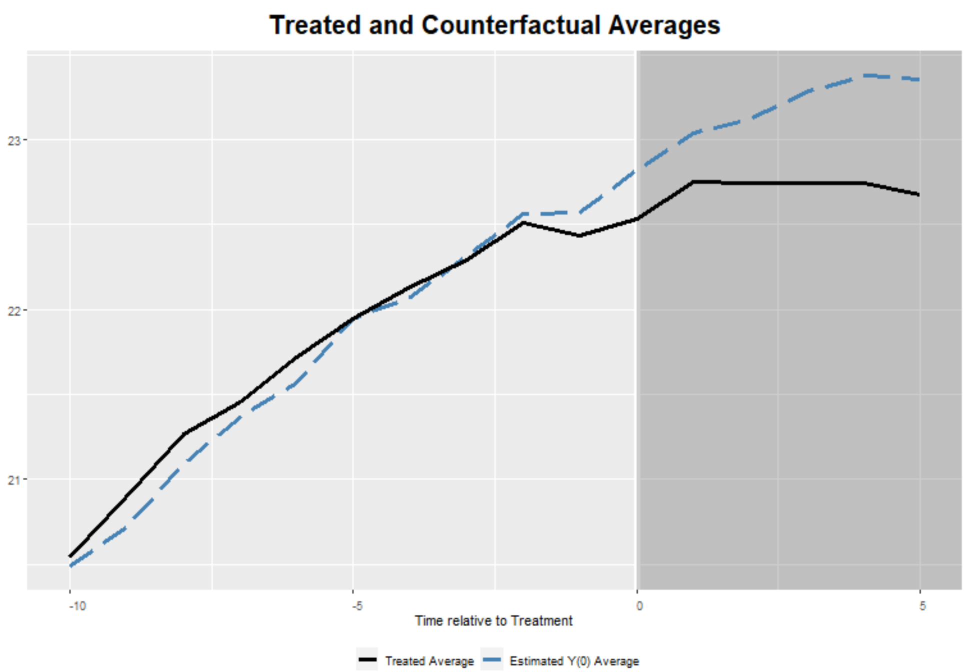


Figure A5: Housing Price Growth, Boom and Bust

Notes: The figure shows the correlation between the housing price growth in 2000-2006 and the one in 2006-2011 across 275 MSAs. The black line is fitted line which with the slope of -1.

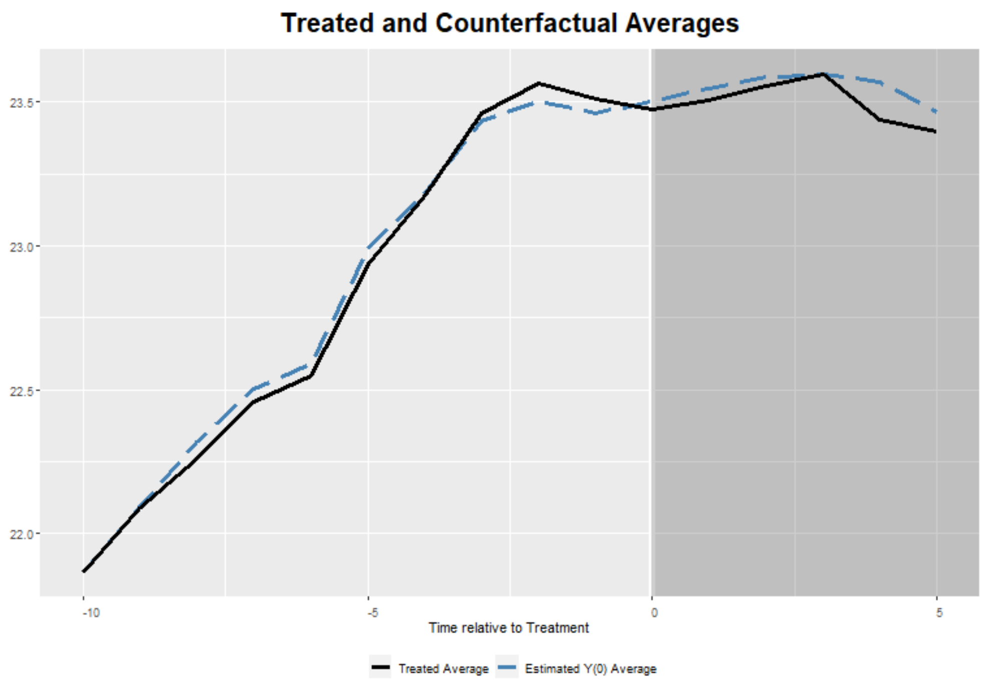


(a) Treated and Counterfactual Averages: Gross Import

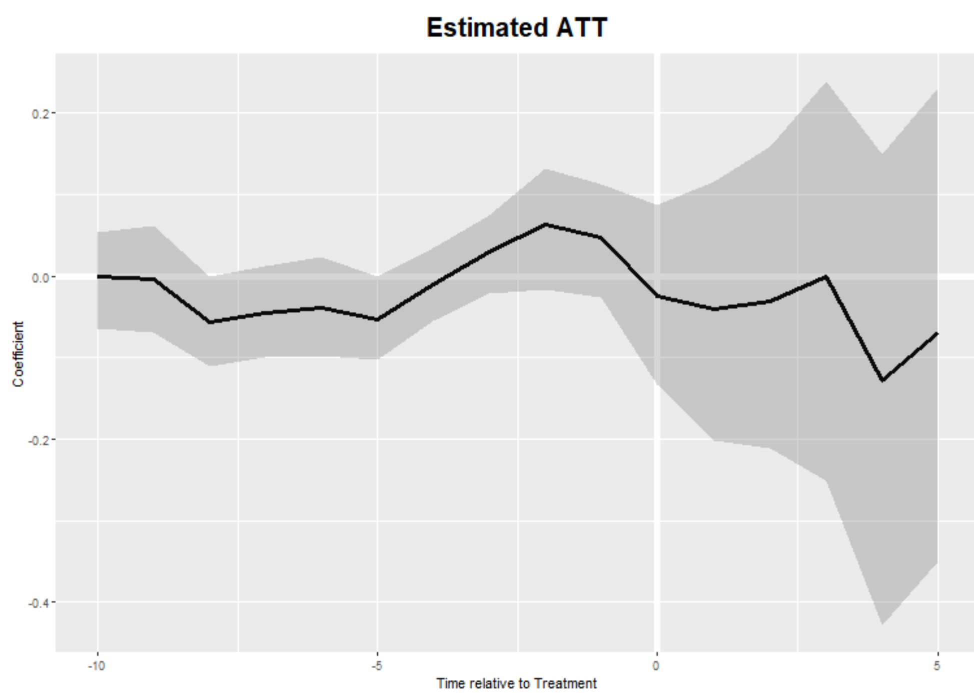


(b) Treated and Counterfactual Averages: Gross Export

Figure A6: The Dynamic Effect of Railways on Trade by DID Method

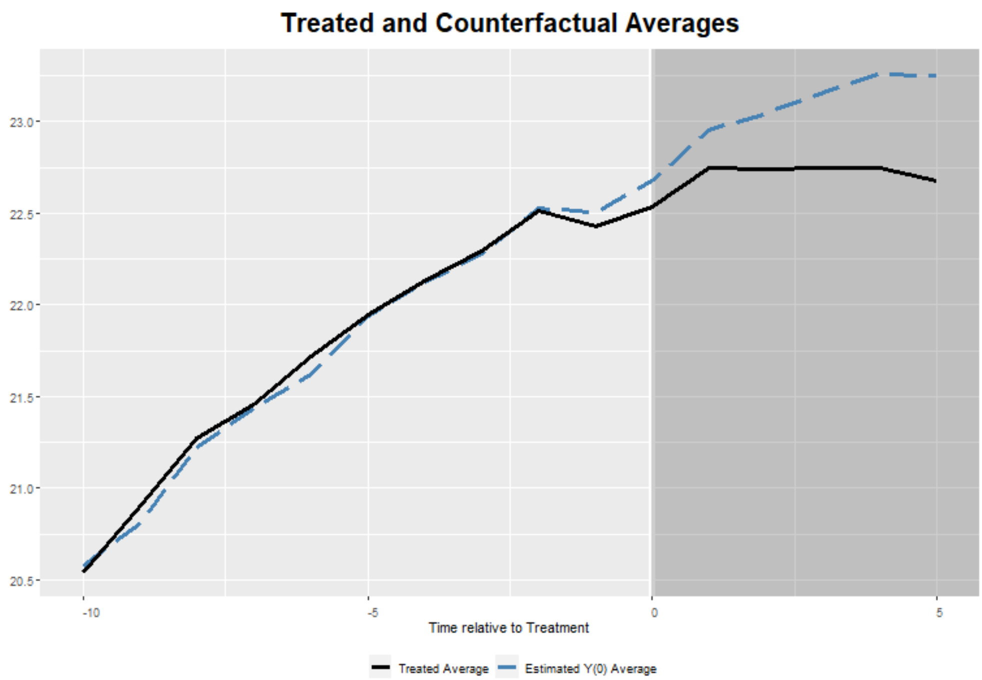


(a) Treated and Counterfactual Averages: Gross Import

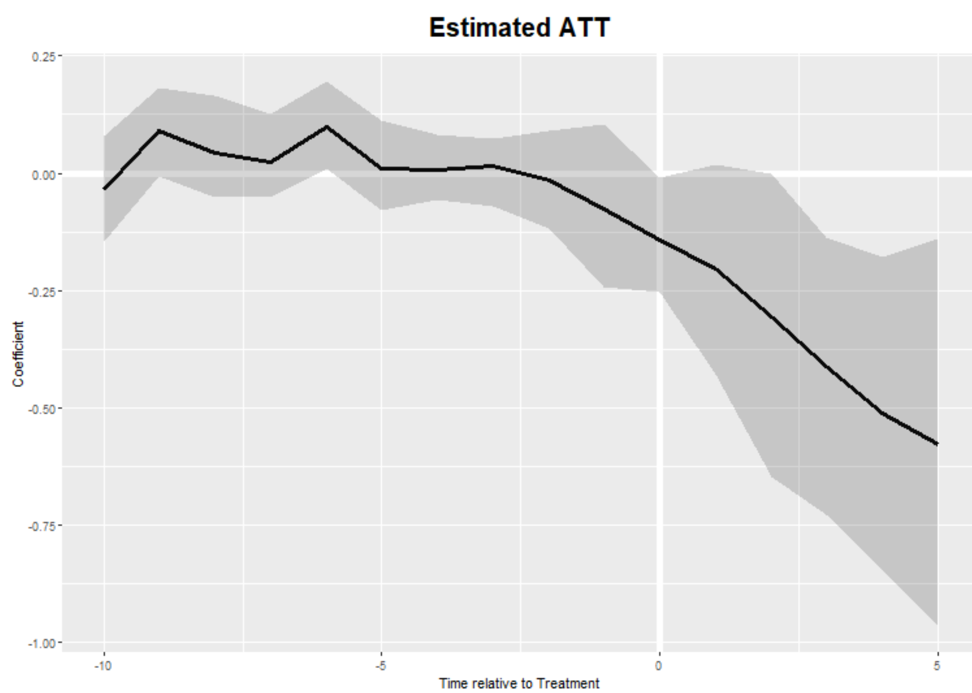


(b) Estimated ATT: Gross Import

Figure A7: The Effect of Railways on Gross Import by Matrix Completion Method

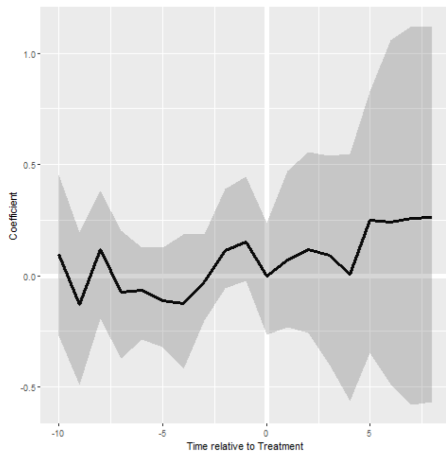


(a) Treated and Counterfactual Averages: Gross Export

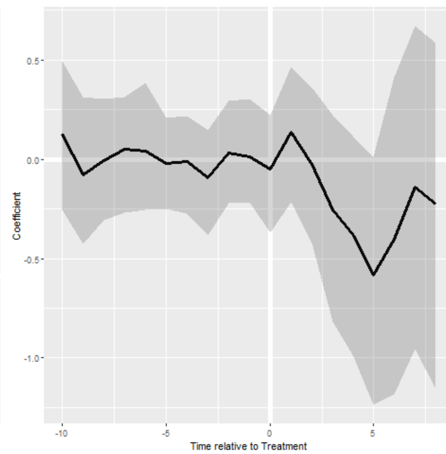


(b) Estimated ATT: Gross Export

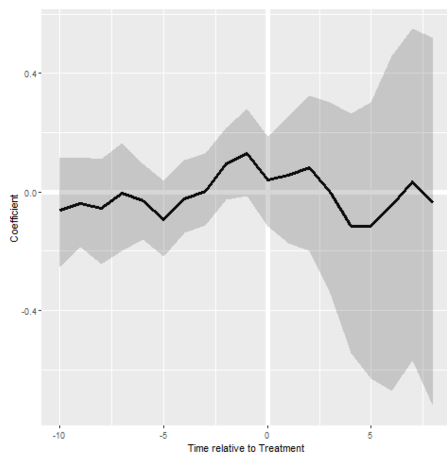
Figure A8: The Effect of Railways on Gross Export by Matrix Completion Method



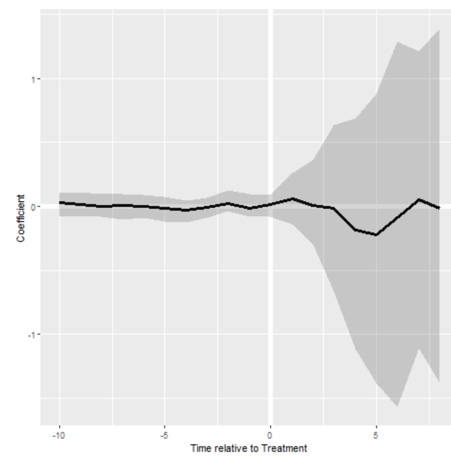
(a) Food and Live Animals



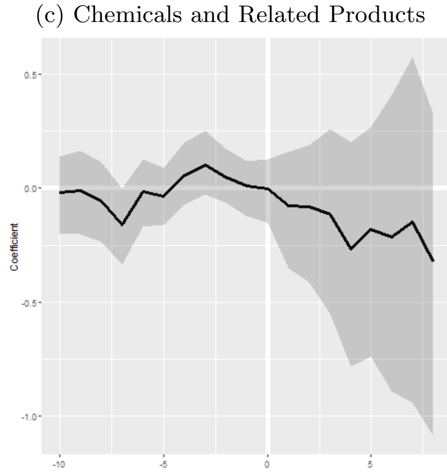
(b) Crude Materials



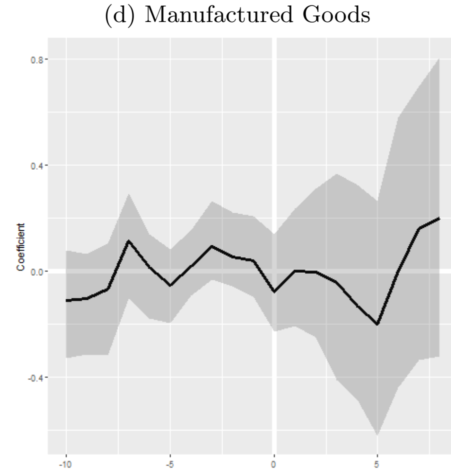
(c) Chemicals and Related Products



(d) Manufactured Goods



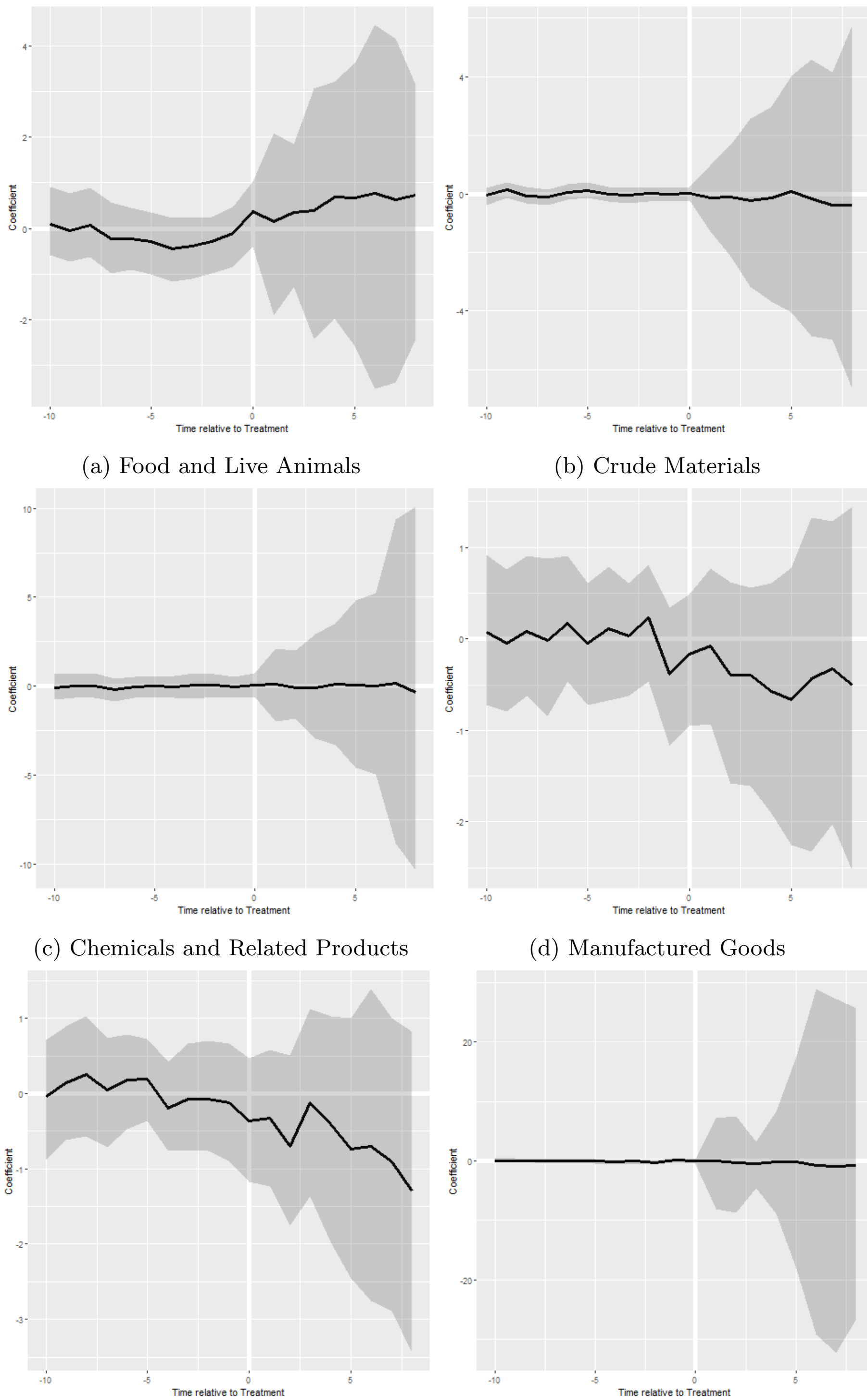
(e) Machinery/Transport Goods



(f) Miscellaneous Manufactured Articles

Figure A9: The Impact of Railways on Gross Import by Product Classifications

Notes: The figure presents GSC estimates of the impact of New Silk Road Railways on trade. The confidence intervals are generated by block bootstrap of 1,000 times.



(a) Food and Live Animals

(b) Crude Materials

(c) Chemicals and Related Products

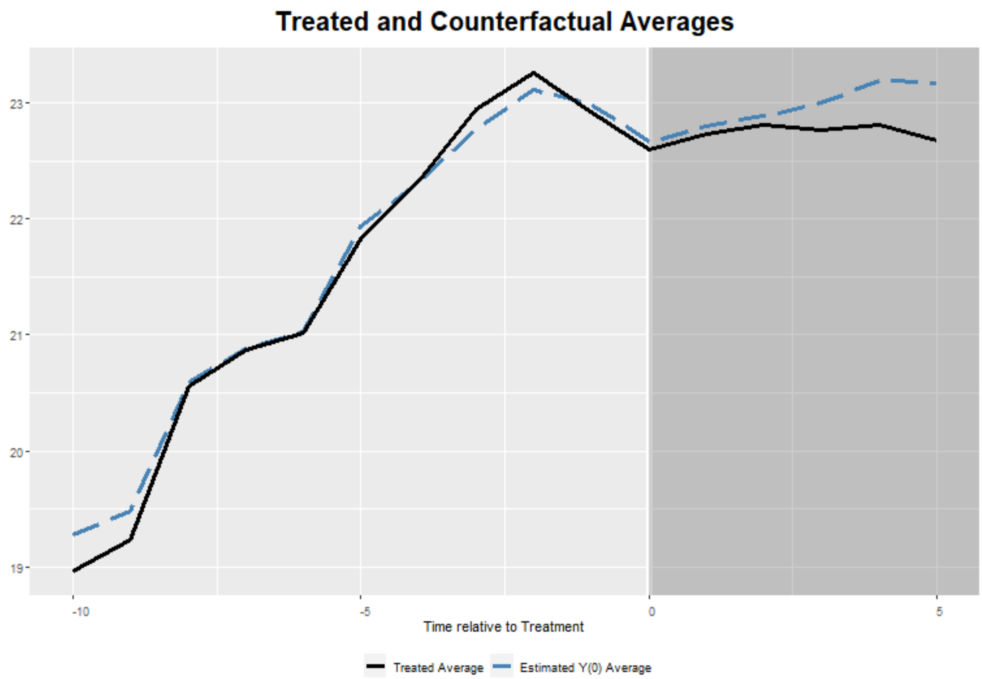
(d) Manufactured Goods

(e) Machinery/Transport Goods

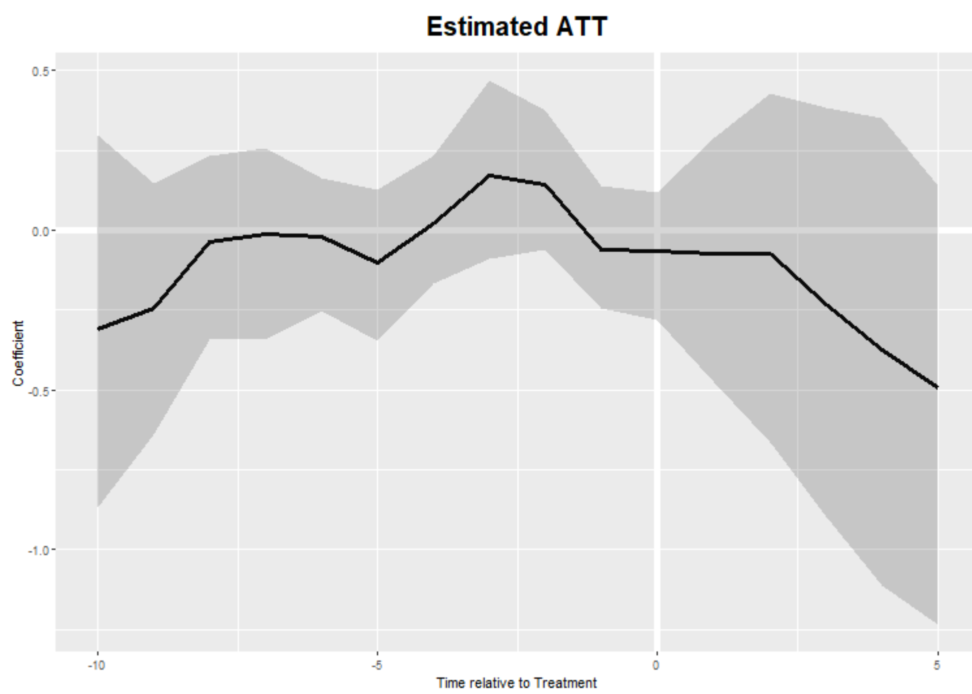
(f) Miscellaneous Manufactured Articles

Figure A10: The Impact of Railways on Gross Export by Product Classifications

Notes: The figure presents GSC estimates of the impact of New Silk Road Railways on trade. The confidence intervals are generated by block bootstrap of 1,000 times.

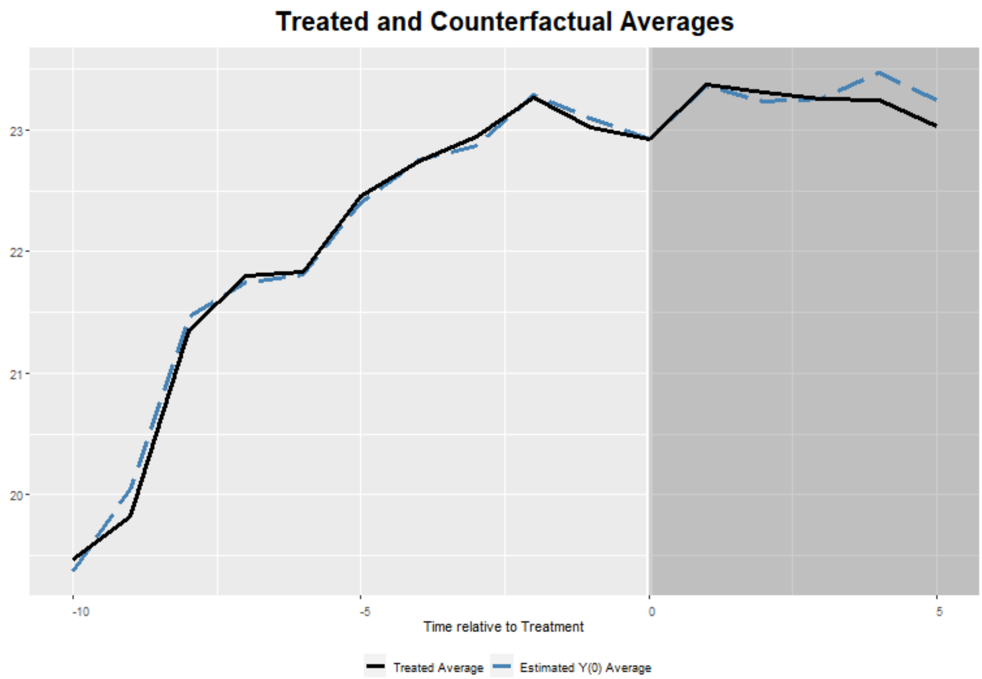


(a) Treated and Counterfactual Averages: Gross Import

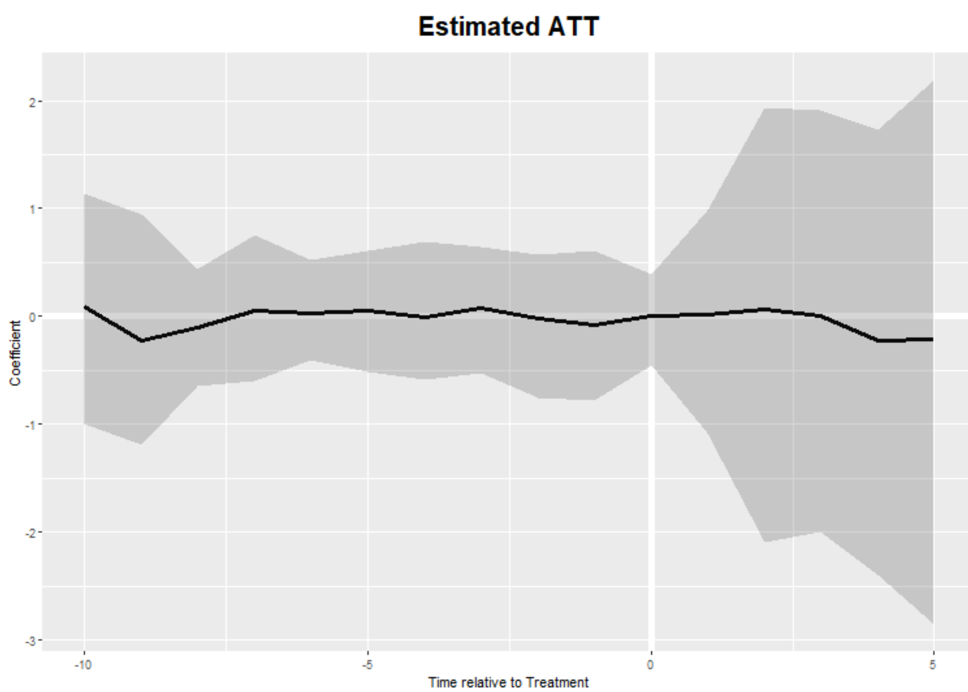


(b) Estimated ATT: Gross Import

Figure A11: The Effect of Railways on Gross Import of Countries Sharing Boundary with China by GSCM

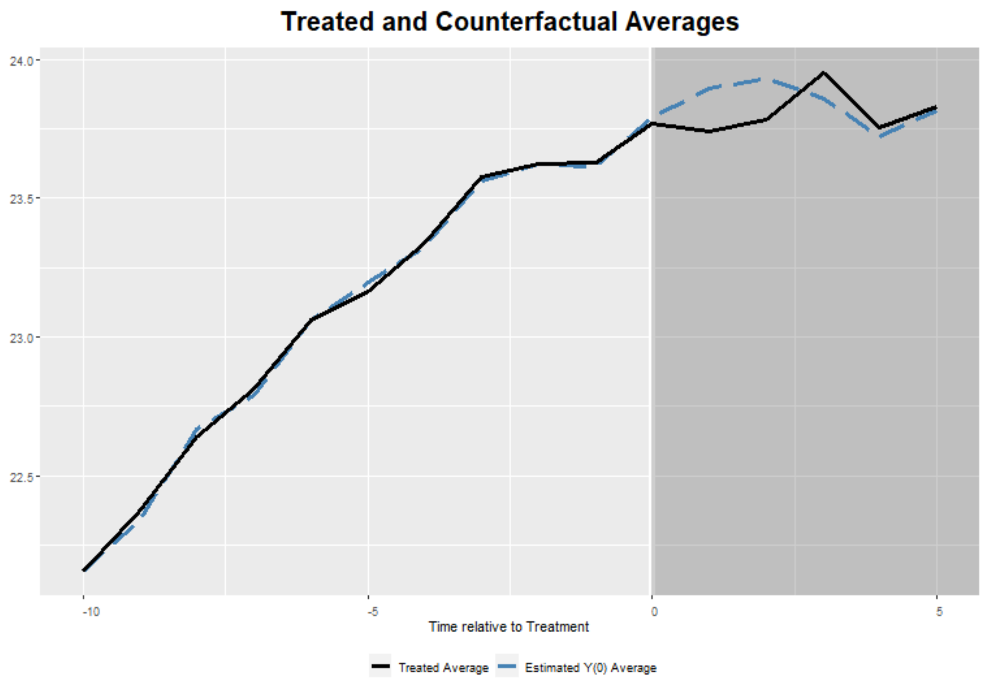


(a) Treated and Counterfactual Averages: Gross Export

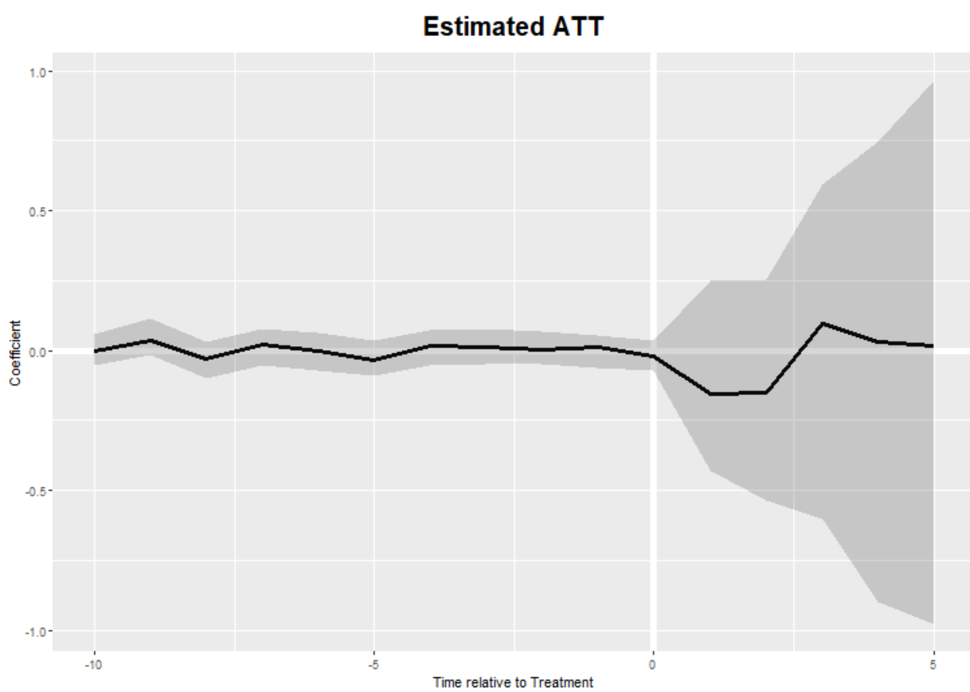


(b) Estimated ATT: Gross Export

Figure A12: The Effect of Railways on Gross Export of Countries Sharing Boundary with China by GSCM

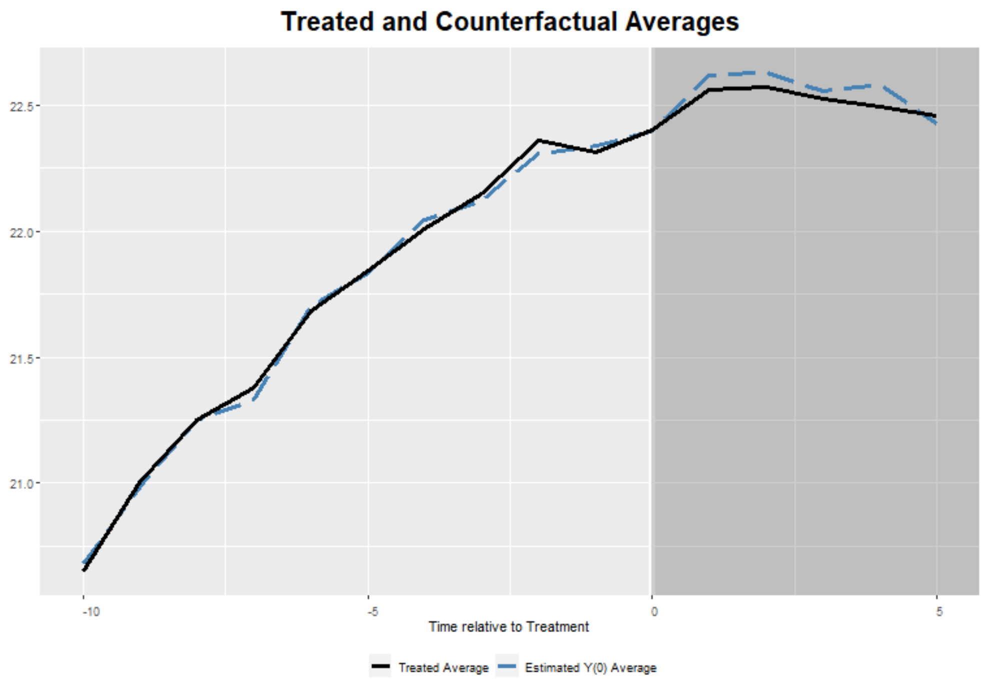


(a) Treated and Counterfactual Averages: Gross Import

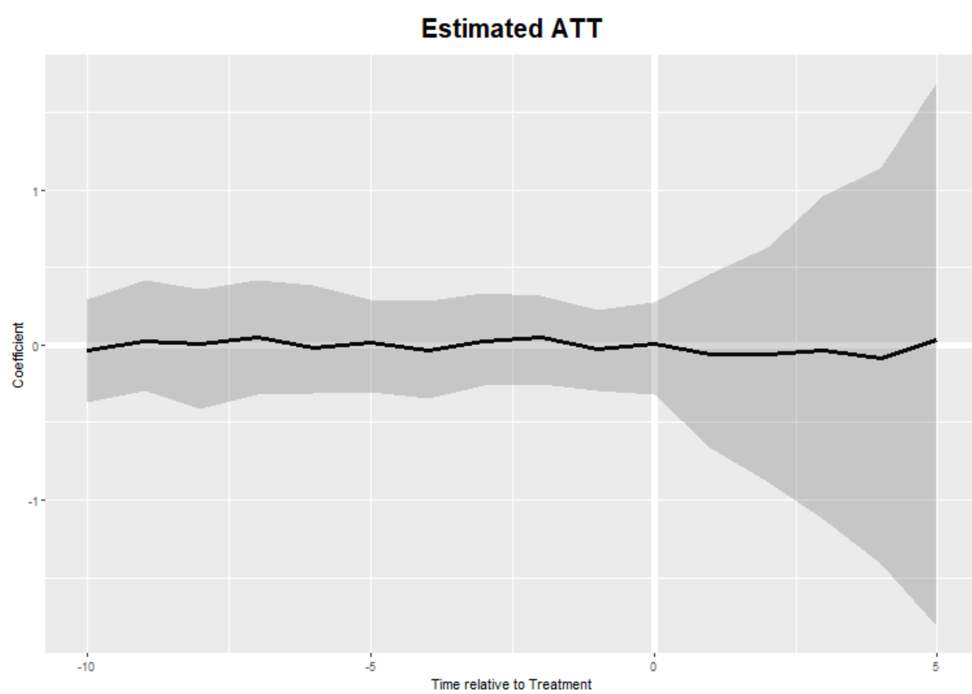


(b) Estimated ATT: Gross Import

Figure A13: The Effect of Railways on Gross Import of Countries not Sharing Boundary with China by GSCM

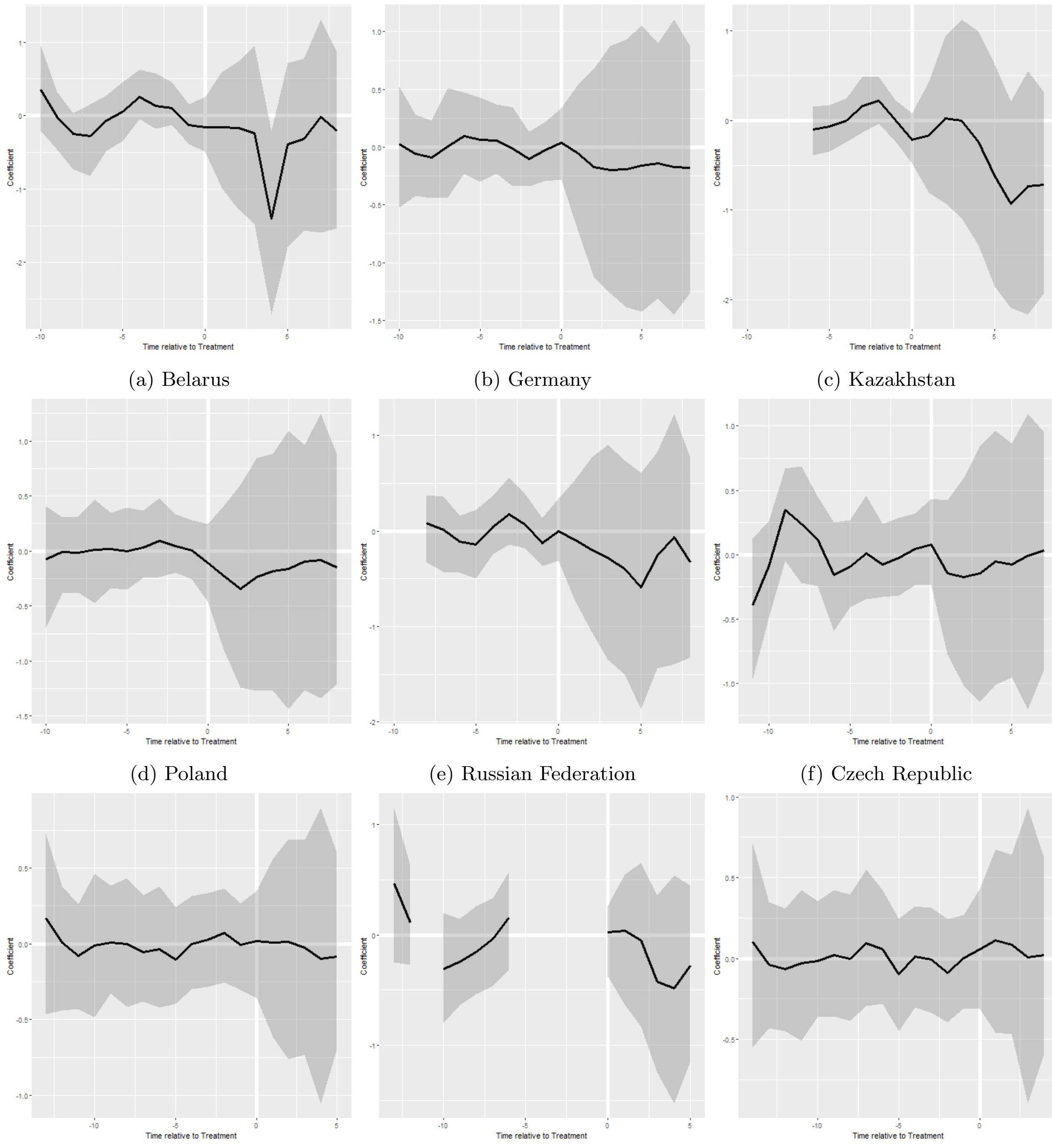


(a) Treated and Counterfactual Averages: Gross Export



(b) Estimated ATT: Gross Export

Figure A14: The Effect of Railways on Gross Export of Countries not Sharing Boundary with China by GSCM



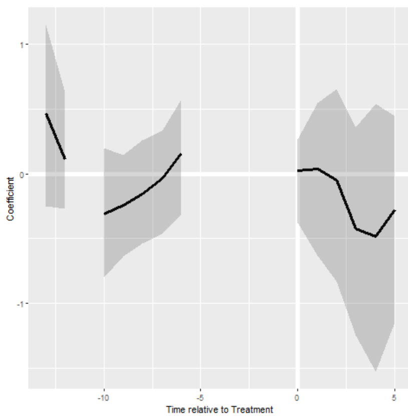
(a) Belarus (b) Germany (c) Kazakhstan

(d) Poland (e) Russian Federation (f) Czech Republic

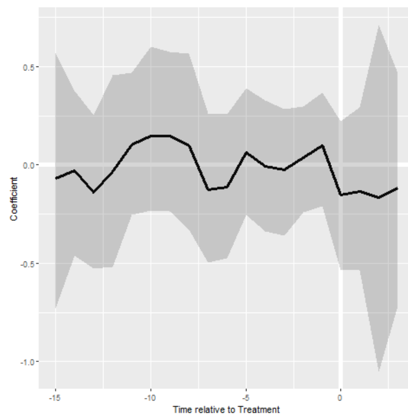
(g) France (h) Mongolia (i) Spain

Figure A15: The Impact of Railways on Gross Import by Product Classifications

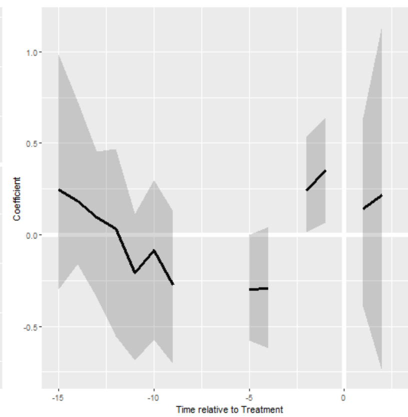
Notes: The figure presents GSC estimates of the impact of New Silk Road Railways on each countries' trade with China. The confidence intervals are generated by block bootstrap of 1,000 times.



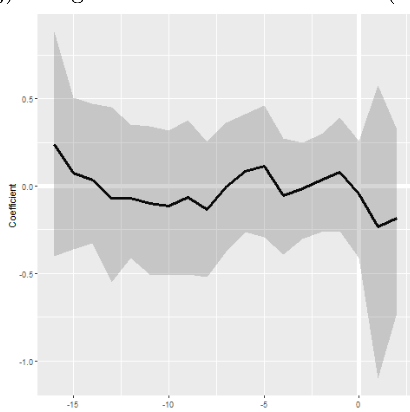
(j) Mongolia



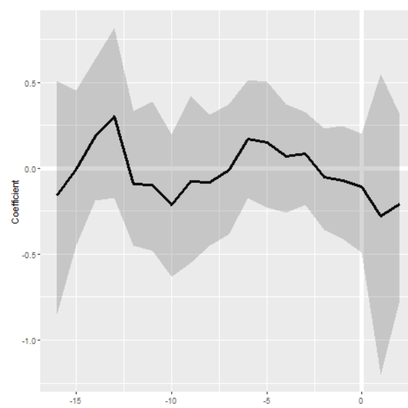
(k) Netherlands



(l) Iran



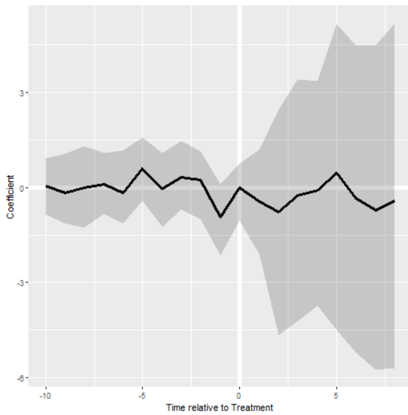
(m) United Kingdom



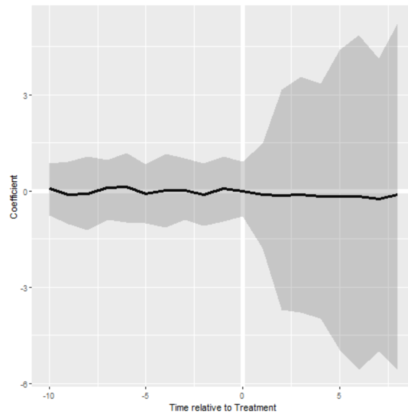
(n) Hungary

Figure A15: The Impact of Railways on Gross Import by Product Classifications (Cont.)

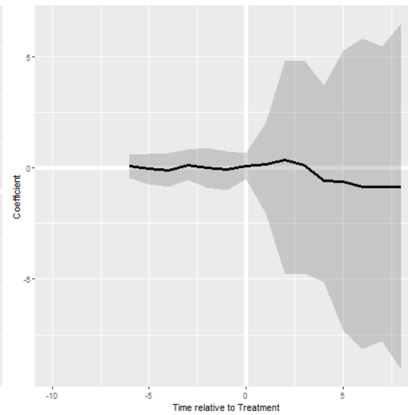
Notes: The figure presents GSC estimates of the impact of New Silk Road Railways on each countries' trade with China. The confidence intervals are generated by block bootstrap of 1,000 times.



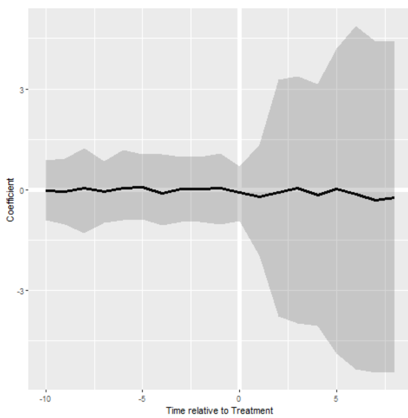
(a) Belarus



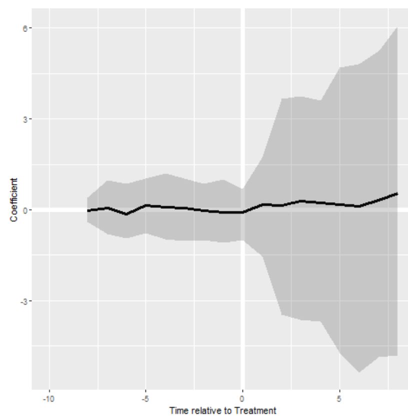
(b) Germany



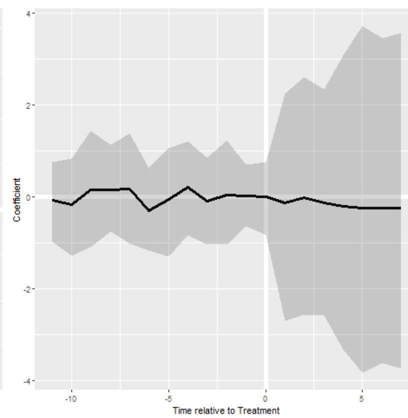
(c) Kazakhstan



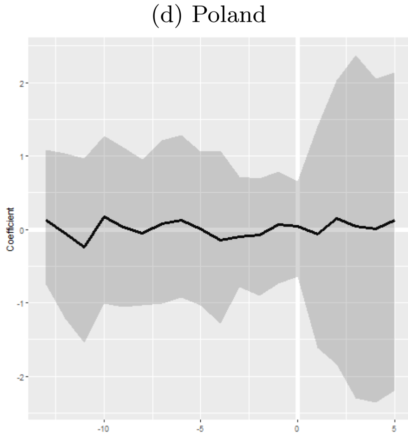
(d) Poland



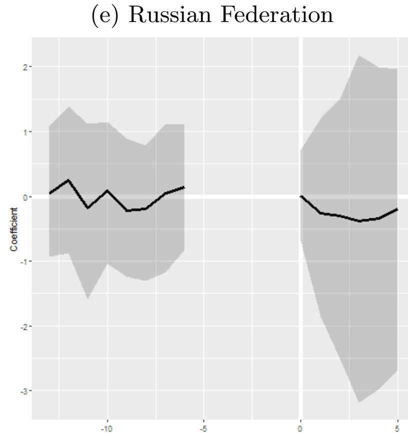
(e) Russian Federation



(f) Czech Republic



(g) France

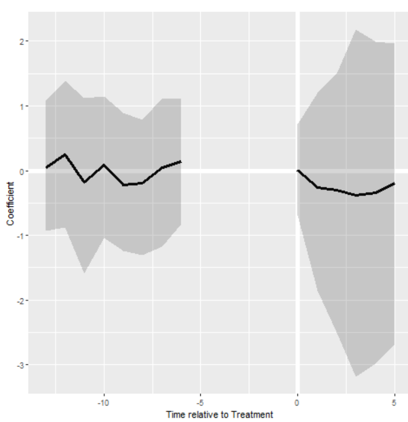


(h) Mongolia

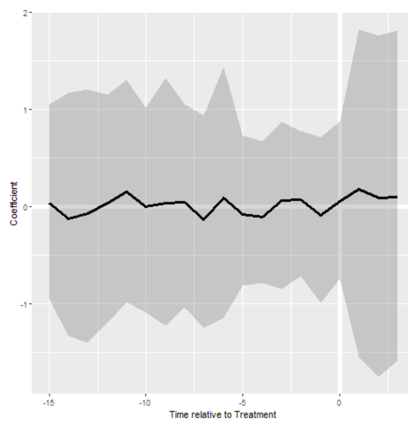
(i) Spain

Figure A16: The Impact of Railways on Gross Export by Product Classifications

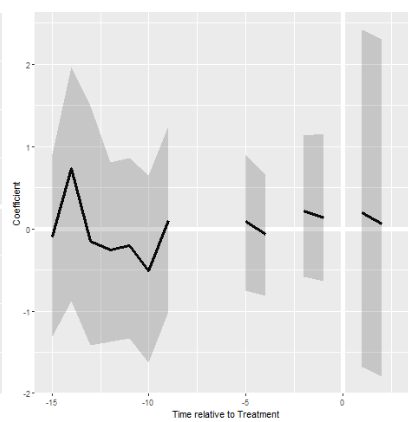
Notes: The figure presents GSC estimates of the impact of New Silk Road Railways on each countries' trade with China. The confidence intervals are generated by block bootstrap of 1,000 times.



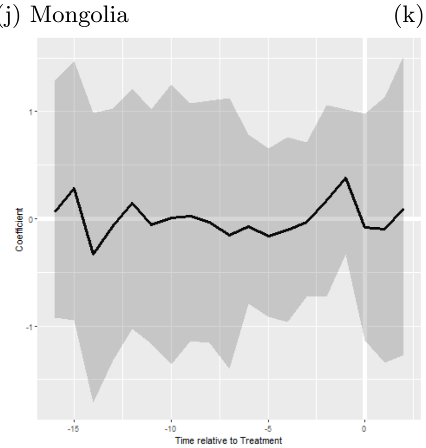
(j) Mongolia



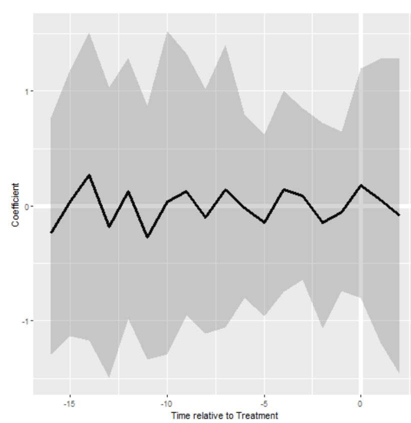
(k) Netherlands



(l) Iran



(m) United Kingdom



(n) Hungary

Figure A16: The Impact of Railways on Gross Export by Product Classifications (Cont.)

Notes: The figure presents GSC estimates of the impact of New Silk Road Railways on each countries' trade with China. The confidence intervals are generated by block bootstrap of 1,000 times.

Appendix Tables

Table A1: Housing Demand Shock and Labor Market Outcomes: OLS Estimates (Full)

Dependent variable: 2000-2006 change in	Emp. Rate	Avg. wage	Emp. Rate	Avg. wage
	Age 18-25		Age 26-33	
Panel A. OLS estimates: all people				
Housing demand change 2000-2006	0.022***	0.064***	0.015***	0.034***
	(0.006)	(0.014)	(0.005)	(0.010)
Log of the age 18-55 population in 2000	-0.005***	-0.005	-0.002*	-0.005
	(0.002)	(0.003)	(0.001)	(0.005)
Female age 18-33 employment rate in 2000	-0.133**	0.084	-0.207***	-0.079
	(0.051)	(0.087)	(0.041)	(0.065)
Employed age 18-33 with college degree in 2000	-0.073	0.040	0.077**	-0.073
	(0.045)	(0.084)	(0.032)	(0.074)
Foreign born share in 2000	0.252	1.356	0.081	2.061*
	(0.540)	(0.880)	(0.408)	(1.075)
Constant	0.130***	-0.052	0.150***	0.159*
	(0.046)	(0.072)	(0.034)	(0.079)
R2	0.331	0.255	0.326	0.243
Panel B. OLS estimates: Men				
Housing demand change 2000-2006	0.025***	0.081***	0.022**	0.042***
	(0.007)	(0.018)	(0.008)	(0.011)
Log of the age 18-55 population in 2000	-0.004**	-0.010**	-0.003	-0.008
	(0.002)	(0.005)	(0.002)	(0.006)
Female age 18-33 employment rate in 2000	-0.139**	0.070	-0.095	-0.146
	(0.058)	(0.110)	(0.078)	(0.088)
Employed age 18-33 with college degree in 2000	-0.081	0.095	0.012	0.021
	(0.055)	(0.099)	(0.040)	(0.092)
Foreign born share in 2000	0.310	1.675*	0.933	1.931
	(0.677)	(0.980)	(0.627)	(1.315)
Constant	0.119**	-0.010	0.079	0.204**
	(0.054)	(0.084)	(0.064)	(0.097)
R2	0.235	0.267	0.206	0.201
Panel C. OLS estimates: Women				
Housing demand change 2000-2006	0.033***	0.040***	0.004	0.018
	(0.008)	(0.010)	(0.008)	(0.011)
Log of the age 18-55 population in 2000	-0.011***	0.001	-0.005*	-0.001
	(0.004)	(0.004)	(0.003)	(0.004)
Female age 18-33 employment rate in 2000	-0.039	0.115	-0.213***	-0.001
	(0.127)	(0.096)	(0.073)	(0.063)
Employed age 18-33 with college degree in 2000	-0.030	-0.063	0.197***	-0.216***
	(0.071)	(0.086)	(0.058)	(0.062)
Foreign born share in 2000	-1.508	0.422	1.807	1.651
	(1.240)	(1.145)	(1.552)	(1.024)
Constant	0.146	-0.091	0.157***	0.108
	(0.106)	(0.096)	(0.055)	(0.083)
R2	0.104	0.067	0.041	0.144
Obs.	275	275	275	275

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A2: Decomposition of Housing Demand Shock and Labor Market Outcomes: OLS Estimates

Dependent variable: 2000-2006 change in	Emp. Rate	Avg. wage	Emp. Rate	Avg. wage
	People Age 18-25		People Age 26-33	
OLS estimates				
All people				
Change in housing prices	0.022***	0.108***	0.023***	0.049**
	(0.008)	(0.021)	(0.008)	(0.021)
Change in housing permits	0.021**	0.030*	0.009	0.023
	(0.009)	(0.015)	(0.005)	(0.015)
P-value of test: $\Delta P_m = \Delta Q_m$	0.944	0.002	0.107	0.374
R2	0.331	0.291	0.332	0.249
Men				
Change in housing prices	0.024**	0.131***	0.013	0.061**
	(0.009)	(0.027)	(0.010)	(0.024)
Change in housing permits	0.026**	0.041**	0.029***	0.028*
	(0.011)	(0.019)	(0.009)	(0.016)
P-value of test: $\Delta P_m = \Delta Q_m$	0.894	0.004	0.117	0.316
R2	0.235	0.298	0.211	0.207
Women				
Change in housing prices	0.057***	0.071***	0.006	0.022
	(0.017)	(0.017)	(0.034)	(0.015)
Change in housing permits	0.015	0.015	0.002	0.014
	(0.015)	(0.014)	(0.018)	(0.018)
P-value of test: $\Delta P_m = \Delta Q_m$	0.122	0.027	0.9245	0.765
R2	0.113	0.081	0.041	0.145
Obs.	275	275	275	275
Baseline controls	Yes	Yes	Yes	Yes

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A3: Housing Demand Shock and Marriage Outcomes: OLS Estimates (Full)

Dependent variable: 2000-2006 change in	marriage rate	
	Age 18-25	Age 26-33
Panel A. OLS estimates: all people		
Housing demand change 2000-2006	0.009***	0.000
	(0.003)	(0.005)
Log of the age 18-55 population in 2000	-0.004**	0.000
	(0.002)	(0.001)
Female age 18-33 employment rate in 2000	0.039	0.131***
	(0.026)	(0.030)
Employed age 18-33 with college degree in 2000	0.142***	0.017
	(0.026)	(0.038)
Foreign born share in 2000	-0.235	0.884
	(0.293)	(0.647)
Constant	-0.052*	-0.167***
	(0.030)	(0.030)
R2	0.166	0.074
Panel B. OLS estimates: Men		
Housing demand change 2000-2006	0.012***	-0.001
	(0.003)	(0.005)
Log of the age 18-55 population in 2000	-0.004*	-0.001
	(0.002)	(0.002)
Female age 18-33 employment rate in 2000	0.051	0.103**
	(0.033)	(0.045)
Employed age 18-33 with college degree in 2000	0.133***	0.026
	(0.028)	(0.039)
Foreign born share in 2000	-0.272	0.813
	(0.303)	(0.682)
Constant	-0.051	-0.122**
	(0.034)	(0.048)
R2	0.163	0.041
Panel C. OLS estimates: Women		
Housing demand change 2000-2006	0.005	0.002
	(0.004)	(0.006)
Log of the age 18-55 population in 2000	-0.003**	0.002
	(0.002)	(0.002)
Female age 18-33 employment rate in 2000	0.022	0.160***
	(0.031)	(0.035)
Employed age 18-33 with college degree in 2000	0.145***	0.006
	(0.026)	(0.049)
Foreign born share in 2000	-0.117	0.920
	(0.404)	(0.738)
Constant	-0.052*	-0.212***
	(0.031)	(0.031)
R2	0.104	0.068
Obs.	275	275

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A4: The 2000-2006 Housing Boom and Marriage Outcomes by Home Ownership

Dependent variable: change in	Marriage rate			
	(1)	(2)	(3)	(4)
Panel A. 2SLS estimates	All Owners	Owners with loan	Owners without loan	Renters
All people age 18-25				
Housing demand change 2000-2006	0.009***	0.009**	0.001	0.006
	(0.004)	(0.004)	(0.001)	-0.009
Men age 18-25				
Housing demand change 2000-2006	0.007**	0.007**	0.000	0.010
	(0.003)	(0.003)	(0.001)	(0.007)
Women age 18-25				
Housing demand change 2000-2006	0.011***	0.007**	0.002	0.002
	(0.004)	(0.003)	(0.001)	(0.011)
Panel B. 2SLS estimates				
All people age 26-33				
Housing demand change 2000-2006	0.002	0.001	0.002	0.002
	(0.009)	(0.010)	(0.002)	-0.011
Men age 26-33				
Housing demand change 2000-2006	0.000	0.001	-0.001	0.003
	(0.011)	(0.012)	(0.002)	(0.013)
Women age 26-33				
Housing demand change 2000-2006	0.006	0.001	0.004*	0.003
	(0.004)	(0.011)	(0.002)	(0.010)
First stage F-statistic	48.940	48.940	48.940	48.940
Obs.	275	275	275	275
Baseline controls	Yes	Yes	Yes	Yes

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A5: Housing Booms and Fertility Rate Defined by Conception

Dependent variable: 2000-2006 change in	Fertility rate	Nonmarital birth rate
Panel A. 2SLS estimates for women age 18-25		
Housing demand change 2000-2006	0.001	-0.038
	(0.008)	(0.039)
Panel B. 2SLS estimates for women age 26-33		
Housing demand change 2000-2006	0.004	-0.022
	(0.006)	(0.018)
First stage F-statistic	46.055	46.055
Obs.	235	235
Baseline controls	Yes	Yes

Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A6: Multinomial Logit Model with Fixed Effects for Fuel Choice, Rural and Urban Households

	Rural		Urban	
	(1) ***Clean Only***	(2) ***Mixed***	(3) ***Clean Only***	(4) ***Mixed***
Household size	0.030	0.127***	0.048	0.107***
	(0.025)	(0.013)	(0.029)	(0.025)
Deopendency ratio	-0.158	-0.178***	-0.148	-0.069
	(0.100)	(0.057)	(0.101)	(0.092)
Main income source: salary	0.918***	0.393***	0.098	0.011
	(0.153)	(0.098)	(0.215)	(0.192)
Main income source: non-agriculture wage	-0.293	-0.222**	-0.257	-0.056
	(0.183)	(0.097)	(0.198)	(0.170)
Main income source: trade	0.500***	0.417***	-0.167	-0.167
	(0.183)	(0.118)	(0.216)	(0.195)
Log real per capita consumption expenditure	0.536***	0.416***	0.729***	0.664***
	(0.094)	(0.058)	(0.119)	(0.113)
Female head	-0.090	0.017	0.201	0.026
	(0.201)	(0.123)	(0.222)	(0.193)
Age of head	-0.001	0.004	-0.002	0.007
	(0.006)	(0.003)	(0.006)	(0.006)
Head's years of education	0.054***	0.043***	0.086***	0.056***
	(0.019)	(0.012)	(0.022)	(0.021)
Log firewood price	0.299***	0.087**	1.106***	0.272
	(0.069)	(0.039)	(0.210)	(0.190)
Log coal price	0.097**	0.019	0.049	0.099**
	(0.045)	(0.025)	(0.045)	(0.047)
Log dung price	-0.086*	-0.203***		
	(0.045)	(0.031)		
Log kerosene price	0.502***	0.431***	0.065	-0.107
	(0.169)	(0.103)	(0.357)	(0.321)
Log LPG price	0.311	-0.215	1.900	-0.329
	(0.451)	(0.186)	(1.443)	(1.291)
Nearest town distance (km)	0.004	0.001		
	(0.007)	(0.004)		
Village has no paved road	-0.039	-0.150		
	(0.162)	(0.092)		
Percentage of households electrified	-0.005*	-0.003**		
	(0.003)	(0.001)		
Village main fuel is clean	0.706***	0.542***		
	(0.158)	(0.107)		
Village main fuel is biomass	0.139	0.375***		
	(0.134)	(0.084)		
Year controls	Yes		Yes	
Log likelihood	-2753.62		-1,626.95	
Observations	10,628	10,628	5,738	5,738

Biomass fuel only is base category. Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A7: Multinomial Logit Model with Random Effects Including Women Bargaining Power Measures

	Rural		Urban	
	(1)	(2)	(3)	(4)
	Clean only	***Mixed***	***Clean only***	***Mixed***
SC/ST	-1.382***	-1.268***	-2.002***	-0.976***
	(0.115)	(0.071)	(0.162)	(0.142)
Muslim	-0.003	-0.070	-1.570***	-0.755***
	(0.145)	(0.090)	(0.169)	(0.148)
Other Backward Class (OBC)	-0.670***	-0.598***	-1.252***	-0.752***
	(0.097)	(0.062)	(0.146)	(0.132)
Household size	0.030*	0.116***	0.011	0.133***
	(0.017)	(0.009)	(0.021)	(0.018)
Deopendency ratio	-0.184**	-0.326***	-0.207**	-0.210***
	(0.077)	(0.045)	(0.085)	(0.073)
Main income source: salary	2.083***	1.178***	2.322***	1.044***
	(0.105)	(0.069)	(0.167)	(0.140)
Main income source: non-agriculture wage	0.327***	0.016	0.402***	-0.093
	(0.121)	(0.063)	(0.156)	(0.125)
Main income source: trade	2.025***	0.848***	1.770***	0.460***
	(0.111)	(0.074)	(0.162)	(0.135)
Log real per capita consumption expenditure	1.661***	1.262***	2.128***	1.226***
	(0.071)	(0.045)	(0.102)	(0.092)
Female head	0.650***	0.585***	1.008***	0.541***
	(0.167)	(0.100)	(0.183)	(0.160)
Age of head	0.036***	0.031***	0.034***	0.020***
	(0.003)	(0.002)	(0.004)	(0.004)
Head's years of education	0.265***	0.187***	0.326***	0.168***
	(0.011)	(0.007)	(0.014)	(0.013)
Log firewood price	0.186***	0.050*	1.187***	-0.072
	(0.051)	(0.030)	(0.143)	(0.122)
Log coal price	-0.041	0.110***	-0.060*	0.290***
	(0.034)	(0.020)	(0.036)	(0.032)
Log dung price	-0.142***	-0.345***		
	(0.037)	(0.024)		
Log kerosene price	0.654***	0.556***	1.632***	0.338*
	(0.123)	(0.078)	(0.210)	(0.180)
Log LPG price	-0.071	-0.258**	-4.200***	-2.682***
	(0.274)	(0.131)	(0.922)	(0.819)
Nearest town distance (km)	-0.009**	-0.011***		
	(0.004)	(0.002)		
Village has no paved road	-0.731***	-0.363***		
	(0.111)	(0.059)		
Percentage of households electrified	0.018***	0.019***		
	(0.002)	(0.001)		
Village main fuel is clean	1.223***	0.733***		
	(0.124)	(0.083)		
Village main fuel is biomass	-0.533***	-0.311***		
	(0.113)	(0.070)		

Biomass fuel only is base category. Standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A7: Multinomial Logit Model with Random Effects Including Women Bargaining Power Measures (Cont.)

	Rural		Urban	
	(1)	(2)	(3)	(4)
	Clean only	***Mixed***	***Clean only***	***Mixed***
Education gap	0.099***	0.070***	0.124***	0.063***
	(0.008)	(0.005)	(0.010)	(0.009)
Violence	0.031	-0.124***	0.078*	-0.032
	(0.037)	(0.023)	(0.044)	(0.039)
Most say in decition macking	0.032	0.022	-0.060	-0.105***
	(0.039)	(0.024)	(0.044)	(0.039)
Financial independence	0.220***	0.195***	0.482***	0.212***
	(0.036)	(0.023)	(0.046)	(0.042)
Freedom	0.025	0.005	0.081*	0.030
	(0.037)	(0.023)	(0.043)	(0.038)
Constant	-20.689***	-13.010***	3.040	4.676
	(1.706)	(0.855)	(5.308)	(4.718)
Heterogeneity covariance				
var(a1)	5.014		6.469	
	(0.452)		(0.607)	
var(a2)	2.837		3.039	
	(0.196)		(0.361)	
Covariance(a1,a2)	2.972***		3.212***	
	(0.258)		(0.415)	
Region controls	Yes		Yes	
Year controls	Yes		Yes	
Log likelihood	-17908.13		-10,600.323	
Observations	36,074	36,074	15,455	15,455

Biomass fuel only is base category. Standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table A8: Multinomial Logit Model with Fixed Effects Including Women Bargaining Power Measures

	Rural		Urban	
	(1) ***Clean Only***	(2) ***Mixed***	(3) ***Clean Only***	(4) ***Mixed***
Household size	-0.025	0.092***	0.027	0.096**
	(0.037)	(0.018)	(0.046)	(0.040)
Deopendency ratio	-0.081	-0.188**	-0.221	-0.124
	(0.158)	(0.091)	(0.164)	(0.144)
Main income source: salary	0.778***	0.139	0.178	-0.020
	(0.239)	(0.145)	(0.331)	(0.290)
Main income source: non-agriculture wage	-0.487*	-0.199	-0.212	-0.254
	(0.263)	(0.137)	(0.292)	(0.244)
Main income source: trade	0.092	0.193	-0.266	-0.215
	(0.255)	(0.167)	(0.324)	(0.290)
Log real per capita consumption expenditure	0.518***	0.290***	1.004***	0.842***
	(0.140)	(0.082)	(0.188)	(0.179)
Female head	-0.788**	-0.208	0.791*	0.476
	(0.393)	(0.228)	(0.407)	(0.341)
Age of head	0.005	0.012**	-0.008	0.005
	(0.009)	(0.005)	(0.011)	(0.009)
Head's years of education	0.029	0.038**	0.090***	0.059*
	(0.029)	(0.018)	(0.033)	(0.033)
Log firewood price	0.286***	0.133**	1.045***	0.574**
	(0.101)	(0.055)	(0.309)	(0.267)
Log coal price	0.161**	0.000	0.065	0.125*
	(0.068)	(0.036)	(0.066)	(0.065)
Log dung price	-0.126**	-0.229***		
	(0.064)	(0.043)		
Log kerosene price	0.768***	0.458***	-0.146	-0.518
	(0.242)	(0.148)	(0.514)	(0.464)
Log LPG price	-0.422	-0.138	-1.171	-3.743**
	(0.679)	(0.215)	(2.131)	(1.907)
Nearest town distance (km)	-0.006	0.005		
	(0.010)	(0.005)		
Village has no paved road	-0.183	-0.123		
	(0.231)	(0.132)		
Percentage of households electrified	-0.006	-0.005**		
	(0.004)	(0.002)		
Village main fuel is clean	0.514**	0.377**		
	(0.241)	(0.154)		
Village main fuel is biomass	0.065	0.291**		
	(0.190)	(0.115)		
Education gap	-0.005	-0.008	0.060***	0.049**
	(0.018)	(0.011)	(0.021)	(0.019)
Violence	0.074	-0.021	0.237***	0.067
	(0.068)	(0.040)	(0.075)	(0.072)
Most say in decition macking	-0.140*	0.016	-0.023	-0.021
	(0.075)	(0.045)	(0.077)	(0.071)
Financial independence	0.048	0.160***	0.118	0.055
	(0.065)	(0.042)	(0.081)	(0.078)
Freedom	0.009	0.038	0.086	0.163**
	(0.037)	(0.023)	(0.043)	(0.038)
Year controls	Yes		Yes	
Log likelihood	-1429.066		-871.67	
Observations	5,716	5,716	3,236	3,236

Biomass fuel only is base category. Robust standard errors in parentheses.
*: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

VITA

Ying-Min Kuo

Candidate for the Degree of

Doctor of Philosophy

Dissertation: THREE ESSAYS ON LABOR AND DEVELOPMENT ECONOMICS

Major Field: Economics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Economics at Oklahoma State University, Stillwater, Oklahoma in May, 2020.

Completed the requirements for the Master of Arts in Economics at Soochow University, Taipei, Taiwan in May, 2010.

Completed the requirements for the Bachelor of Arts in Economics at Soochow University, Taipei, Taiwan in May, 2008.

Experience:

Graduate Teaching and Research Assistant
January 2020 - May 2020

Graduate Student Instructor, Introduction to Macroeconomics
August 2018 - December 2019

Graduate Teaching Assistant
August 2015 - May 2018

Professional Memberships:

Fulbright Scholar
August 2015 - May 2020