PHBS WORKING PAPER SERIES

Home Location Choices and the Gender Commute Gap

Yizhen Gu
Peking University

Naijia Guo
Chinese University of Hong Kong

Jing Wu
Tsinghua University

Ben Zou
Michigan State University

January 2021

Working Paper 20210105

Abstract

Using administrative records of home mortgages in Beijing, we show that households systematically choose to buy new homes that are closer to the wife's workplace. Compared with the husband's commute, new homes are on average 11% closer to the wife's workplace by distance and requires 4% less commute time. In terms of a household's disutility from choosing a particular home, a one log point increase in the wife's commute distance is equivalent to a 0.5 log point increase in the house price, and a 1.7 log point increase in the husband's commute distance. Using a collective household model, we show that home location choices and the resulting gender commute differences reflect the intra-household division of labor and differences in bargaining power. We discuss the implications of the observed gender commute gap for the gender pay gap.

Keywords: Gender gap, commute to work, housing demand.
JEL Classification: J16, R21, J22

Peking University HSBC Business School
University Town, Nanshan District
Shenzhen 518055, China

Home Location Choices and the Gender Commute Gap

Yizhen Gu, Naijia Guo, Jing Wu and Ben Zou*

October 25, 2020

Abstract

Using administrative records of home mortgages in Beijing, we show that households systematically choose to buy new homes that are closer to the wife's workplace. Compared with the husband's commute, new homes are on average 11% closer to the wife's workplace by distance and requires 4% less commute time. In terms of a household's disutility from choosing a particular home, a one log point increase in the wife's commute distance is equivalent to a 0.5 log point increase in the house price, and a 1.7 log point increase in the husband's commute distance. Using a collective household model, we show that home location choices and the resulting gender commute differences reflect the intra-household division of labor and differences in bargaining power. We discuss the implications of the observed gender commute gap for the gender pay gap.

JEL classification: J16, R21, J22

Keywords: Gender gap, commute to work, housing demand.

1 Introduction

There are substantial gender differences in various aspects of the labor market. Men and women have different labor force participation rates and working hours; they work in systematically different occupations and industries; and women earn substantially lower wages for the same work. An extensive body of research documents and explains these gender differences (for reviews, see Altonji and Blank, 1999; Marianne, 2011).

There is also a substantial gender difference in the length of the commute to work. In the United States, female workers on average spend 11% less time in commute than male workers.[1] In OECD countries, women's commutes are on average 33% shorter than men's (Organization for Economic Co-operation and Development, 2015, Table LMF2.6). The gap is even larger among married people and

*Gu: Peking University, yizhengu@phbs.pku.edu.cn; Guo: Chinese University of Hong Kong, guonaijia@cuhk.edu.hk; Wu: Tsinghua University, ireswujing@tsinghua.edu.cn; Zou: Michigan State University, benzou@msu.edu.

[1] These numbers are from the authors' calculations using the 2017 American Community Survey, and are based on individuals aged 25 to 64 years old who work at least 20 hours a week. The socioeconomic conditions include a set of education attainment indicators, a set of age group indicators, marital status, whether having children under 6 years old, a set of metropolitan area indicators, and a set of transportation mode indicators. See Appendix Table B.1 for details.

those with children. Although the gender commute gap is evident in many countries and persistent over time, as documented in the urban studies literature (e.g., Madden, 1981; MacDonald, 1999; Crane, 2007), little economic research examines this issue. The few exceptions (e.g., White, 1977,9, and more recently Gutierrez, 2018; Le Barbanchon et al., 2019; Liu and Su, 2020) focus on the potential relationships between the gender commute gap and other gender differences in terms of labor market outcomes.

There are two possible explanations for the observed gender gap in commutes. The first explanation relates to the labor supply and job searching, and proposes that conditional on their location of residence, women systematically search and accept jobs that require shorter commutes. The second relates to the choice of home location, and proposes that conditional on the location of the workplace, women systematically live closer to their places of work. Among married couples, this means that households systematically choose home locations that are closer to the wife's workplace.

Although these two explanations are not mutually exclusive and both could be driven by women's relative disutility from commuting, they have different implications for the gender pay gap. If the gender gap in commuting is driven by women not being willing to accept jobs that require long commutes, shorter commutes among women will be directly linked to lower earnings because conditional on where one lives, searching over a larger radius is more likely to result in finding a higher-paying job. In contrast, a reluctance to commute limits the set of job offers women can choose from. This view is supported by recent studies (e.g., Gutierrez, 2018; Le Barbanchon et al., 2019; Liu and Su, 2020), which shows that between 10% and 20% of the gender wage gap is accounted for by gender differences in the willingness to pay for a shorter commute.

In contrast, if households systematically choose to live closer to the wife's workplace, the observed gender gap will possibly result from a household's collective effort to mitigate the gender differences in the labor market outcomes. Because women traditionally play a greater role in household production, commuting is likely to be particularly costly for them. Living closer to where the wife works makes it easier for her to stay in her job and remain in the labor market when the household moves.

It is difficult to empirically disentangle these two potential explanations largely due to the lack of data. Few datasets contain both commute and labor market information, and those that do are mostly cross-sectional in nature. Cross-sectional data do not allow us to determine whether the observed gender commute gap is due to job search decisions or home location choices.

In this paper, we investigate whether households systematically choose to buy homes that are closer to the wife's workplace than the husband's. To achieve this, we leverage a unique dataset of administrative records of home mortgages in Beijing. When a household applies for a mortgage, detailed information is collected to evaluate the borrower's creditworthiness. Importantly for our purposes, the dataset includes detailed addresses of the current and new homes, and the work addresses of the borrower and the co-borrower (usually the lender's spouse). From these addresses we calculate the couple's commute *distances* from the current and new homes. The dataset also includes information on the borrower and co-borrower's demographic and economic characteristics, including age, education, and annual income. We treat the couple's jobs as given, and test how they choose their new home location given the husband's and wife's work locations.

We first document a substantial and robust gender gap in the commute distance from the *current* home. In a sample of households in which the husband and wife both commute to work, the wife's commute is between 11% and 14% shorter than the husband's. The gender commute gap is also correlated with the gender earnings gap. On average, a female worker in Beijing earns about 25.7% less than a male worker with the same qualifications. The gender earnings gap shrinks to 23.4% when we control for the commute distance. Therefore, the gender commute gap accounts for about 9% of the gender earnings gap, which is similar to the figures found in recent studies from mostly developed countries (Gutierrez, 2018; Le Barbanchon et al., 2019).

We then show that when deciding where to locate their *new* homes, households systematically choose homes that are closer to the wife's workplace. The wife's commute from the new home is on average between 11% and 13% shorter in distance than the husband's. Although the average commutes from new homes are almost 40% longer as households tend to buy new homes that are farther away from the city center, where most jobs are located, the gender commute gap in log terms remains roughly constant. To account for the differential sorting into workplaces and locations of residence, we control for detailed demographic characteristics such as age and education, and a set of fixed effects of the home districts. We also restrict the gender commute gap within a household. The results are essentially unchanged.

Using a discrete choice random utility model, we find that the commute distance plays an important role in a household's decision on which home to buy. In our preferred specification, a one log point increase in the wife's commute distance has a marginal effect on the probability of a particular home being chosen that is equivalent to the marginal effect of a 0.5 log point increase in the house price. The household derives a larger disutility from the wife's commute than from the husband's, with a one log point increase in the wife's commute distance being equivalent to a 1.68 log point increase in the husband's commute distance. Therefore, a household's choice of residence location plays an important role in the observed gender gap in commutes.

We investigate whether our results are sensitive to job switches after the household takes out a mortgage. If households systematically buy new homes that are closer to the husbands' *future* work locations, our results will overstate the gender commute gap from new homes. Using a panel of employer-employee linked records matched with our mortgage data, we show that the gender commute gap remains up to 9 years after a household takes out a mortgage. Workers do switch employers, but there are no statistically discernible differences in the job switch or job loss rates before and after obtaining a mortgage, or between men and women. If anything, the point estimators suggest that after taking out mortgages, the husbands tends to switch to jobs that are farther away from the new home, and the gender commute gap increases.

As people tend to use faster modes of transportation for longer commutes and those commutes thus tend to have a higher average speed (Couture et al., 2018),[2] and, the gender gap in commute *time* may

[2] All commutes involve "fixed costs," such as accessing the car, walking from the parking lot, walking to the subway station or bus stop, and the wait time at the subway station or bus stop, that tend to be diluted over longer commutes. Longer commutes also tend to involve traveling on highways and suburban roads, which tend to have higher road speeds than the

be smaller than that in commute *distance*. The mortgage data do not include information on commute times. To overcome this shortcoming, we take advantage of the exact work and home addresses in the data, and use Baidu Maps — a popular digital map service in China — to obtain the commute times for different modes of transportation during the weekday rush hours. Using the 2015 Beijing Household Travel Survey, we fit a model of transportation mode choices for the husband and wife based on the couple's demographic characteristics, household income, commute distances, and access to public transit. We then impute the mode choice for each individual in the mortgage data and assign the corresponding commute time from Baidu Maps. We find that the gender gap in commute *time* shrinks to about 3.5% but remains statistically significant.

We also use the Household Travel Survey to provide corroborative evidence. The survey provides direct information on commute times. The estimated gender gap in commute *time* from this sample is about 10%. The Household Travel Survey is cross-sectional and does not cover households' choices of home location. However, by focusing on a subset of households that have recently moved, we find a similar, though somewhat smaller gender commute gap of 6%. To account for the amenity differences across different modes of transportation, we further control for a set of transportation mode fixed effects. Because choosing a faster mode of transportation is positively correlated with the length of commute, controlling for mode fixed effects leads to a lower bound of the gender commute gap. Nevertheless, we continue to find a statistically significant and economically meaningful gender commute gap.

The household decision to live closer to the wife's workplace reflects women's relative distaste for commuting. This relative distaste may stem from the traditional gender differences in the household and market production roles. We build a simple model of intra-household bargaining to shed light on the collective household location decision and the gender gap in commutes. The model predicts that the couple will live closer to the wife's workplace when (1) the wife's time efficiency in home production is higher, (2) the value of the couple's public goods is higher, and (3) the wife's bargaining power is higher.

To empirically test these predictions, we explore household characteristics that are associated with the magnitude of the commute gap between the husband and wife. We first show that the wife's commute is shorter than the husband's in virtually all types of households. Consistent with the prediction regarding the household division of labor, we find that the gender commute gaps are larger among married couples, especially those with children. Consistent with the prediction regarding intra-household bargaining power, the gender gap is smaller among households in which the husband's relative education or income is higher.

This paper is related to several strands of the literature in labor and urban economics, and contributes to the small but growing body of research on the causes and implications of the gender commute gap. Economics and urban studies researchers have long noticed the gender commute gap, and suspected that it is related to the gender wage gap (e.g., Madden, 1981; MacDonald, 1999; Crane, 2007). As better data are now available, research interest in this issue has been reignited. Black et al. (2014) find that married women in cities with longer commutes are less likely to be in the labor force. Gutierrez (2018)

urban roads in downtown areas.

documents a substantive and persistent gender commute gap in the United States. He goes on to show that the gender wage gap declines substantially once the commute length is controlled for, suggesting that the two types of gender difference are connected. Using administrative data from France on job search criteria and exploiting a unique institutional feature that requires unemployed workers to report the job offers they receive, Le Barbanchon et al. (2019) show that women have a smaller radius for job searches and are willing to accept lower wages for jobs closer to home. These studies suggest that the gender commute gap and gender wage gap are connected through the labor supply channel. Using the American Community Survey, Liu and Su (2020) take a similar supply-side view and show that the gender commute gap contributes to between 16% and 21% of the gender wage gap in the United States. We see our results as complementing those focusing on the labor supply and job searches. We show that the household decision on the location of the home may also contribute to the observed gender commute gap.

This paper is related to the large body of literature on household bargaining and the gender gaps in labor market outcomes that has been published since the seminal works by Becker (2009) and Chiappori (1988,9). Recent studies suggest that working women are particularly time-constrained due to their traditional role in household production (e.g., Bertrand et al., 2010; Goldin, 2014; Cortés and Pan, 2019). The time spent on commuting to work further tightens these time constraints and increases the costs of working. This paper shows that collective household decisions imply a shorter commute for the wife, and could potentially improve women's labor market participation and reduce the gender wage gap. Moreover, the gender commute gap is larger when women have greater bargaining power and when a couple has children.

Our findings are relevant to two sets of models commonly used in the labor and urban economics literature. First, in the neighborhood choice models, the home location choices are endogenized (Bayer et al., 2007, Bayer et al., 2016). However, these models typically do not consider commuting. Moreover, in the few exceptions in which commuting is considered, typically only one work location is allowed per household (Kuminoff et al., 2012; Li et al., 2018). In contrast, the recent spatial general equilibrium models for cities take commuting seriously (Monte et al., 2018; Tsivanidis, 2018; Severen, 2019). However, they typically model individual rather than collective household decisions. The empirical pattern documented in this paper suggests that it is worth building a general equilibrium model that incorporates both endogenous home location choices and the collective household labor supply, which we leave for future work.

The rest of this paper is organized as follows. Section 2 introduces the data and documents basic empirical facts about commuting in Beijing. Section 3 presents the empirical results on home location choices. Section 4 builds a simple intra-household bargaining model and exploits the potential mechanisms. Section 5 concludes the paper.

2 Data and Sample

2.1 Data

The main data used in this paper are the administrative records of home mortgages issued in Beijing from an anonymous mortgage lender. The lender is a major player in Beijing's housing market. In 2013 for example, the lender accounted for about 15% of total home sales in the city. Our sample comprises 183,005 records of home mortgages between 2005 and 2014. When a household takes out a mortgage, the borrower is required to provide detailed information on her employment status, occupation and industry, monthly salary, current home address, and a host of questions on other sources of income and assets. The mortgage contract is typically co-signed by a co-borrower, usually the main borrower's spouse. The vast majority of the borrowers in our sample are married, thus we also have detailed information on the spouse. We drop mortgages with sole borrowers and those whose co-borrowers are unlikely to be a spouse.[3] Over 87% of the households are first-time home buyers. We leave all households in the sample. The results are unchanged when the sample is restricted to only first-time home buyers.[4]

Important for our purposes, the data contain detailed addresses of the couple's workplaces and the addresses of their current home and new home (the one with a mortgage). In the sample, 94% of men and 84% of women are employed and provide a valid employer address. From these addresses we are able to calculate the linear *distances* from the current and new homes for both the husband and wife. These are our main measures of the commute lengths. We also calculate the road distances and commute *times* under different modes of transportation from Baidu Maps.[5] The mortgage data do not include information on the modes of transportation used for commutes. Thus, we impute the modes of transportation for both the husband and wife based on the couple's demographic characteristics, household income, length of commute, and access to public transit.

To corroborate the evidence from the mortgage data, we also use the 2015 Beijing Household Travel Survey, which covers around 100 thousand individuals from 40 thousand households in Beijing. The survey records where each respondent lives and works,[6] and includes a travel diary that tracks each respondent's whereabouts on the survey day. For each trip, the diary records the departure and arrival locations and times, and the modes of transportation used. From the travel diaries, we compile 62,697 commuting trips (to and from work) for 28,366 workers. The dataset also includes a roster and demographic characteristics of each household member, and household income, although it does not include individual income.

It is worth noting that neither the mortgage dataset nor the travel survey is representative of the population of Beijing. The mortgage data only include households who purchase a home and take out a

[3]Only 93 mortgages have borrowers that are unmarried and only two mortgages do not have a co-borrower. The data do not indicate the relationship between the borrower and the co-borrower. We drop around 3,000 mortgages in which the borrower and the co-borrower are of the same sex, or the age gap is more than 20 years (which may indicate that the borrower and co-borrower are parent and child).

[4]Appendix Table B.2 presents the robustness check.

[5]http://maps.baidu.com/. The developer platforms are available at https://lbsyun.baidu.com/.

[6]The smallest identifiable geographic area in the Household Travel Survey is the Transportation Analysis Zone (TAZ). Beijing, with an area of 6,490 square miles and a population of 21.5 million, is divided into approximately 2,000 TAZs.

mortgage from the anonymous lender. Although intended to be representative of the city's population, the Household Travel Survey suffers from some flaws in terms of design and implementation. Based on voluntary take-up at the housing unit level, the sample over-represents older residents (over half of the sample are retirees) and homeowners. Unfortunately, because we do not have detailed demographic data at geographic units below the county/district level (Beijing was divided into 18 counties/districts), we are unable to construct weights for either sample to restore representativeness.[7]

2.2 Summary Statistics

Beijing has undergone large-scale urbanization and suburbanization over the past two decades. The city had a population of about 21 million in 2019, almost double that of 2000. However, due to the rapidly growing housing prices, families have had to increasingly live farther away from the city center. This is well borne out in the mortgage data, with Panel A of Table 1 showing that new homes are about 10% farther away from the city center than current homes. This is largely because about two thirds of the homes purchased in the sample are more than 15 kilometers away from the downtown, as shown in Panel A of Figure 1. Commutes from new homes are about 40% farther away from the workplaces. From current homes, women's commutes are 11% shorter than men's; from new homes, the gender gap in commutes remains at 9%.

The mortgage data also reveal the spatial distribution of jobs in Beijing. A large portion of Beijing's city center is occupied by museums, parks, and monuments. The employment density is highest between 2.5 and 5 kilometers from the city center (Panel C of Figure 1). Similar to cities around the world, jobs in the city center pay the highest wages. Wages decline as the distance from the city center increases (Panel D). The spatial distributions of jobs occupied by men and women are similar (Panel B). As more households buy homes that are farther away from the city center (Panel A), the commute distances increase for both the husband and wife.

The sample from the 2015 Household Travel Survey is on average older and less educated than the mortgage sample, and it also includes a substantially greater proportion of unmarried individuals. The commuting distance, measured as the linear distance between the centroid of the home TAZ and that of the work TAZ, is on average 6.36 km for men and 5.02 km for women. The average commuting distances in the survey sample are also much shorter than those in the mortgage sample, because the survey is over representative of the TAZs in the central districts. The gender gap in commute *distance* is about 21%. The gender gap in commute *time* is about 7%, with an average one-way commute time of 34.7 minutes for men and 32.2 minutes for women. The smaller gap in commute *time* is probably due to the fact men are more likely to drive, with 34% of men driving to work, compared with 19% of women. About 10% of men and women take the subway to work. Finally, women are more likely to commute by slower modes such as bus, bike, and walking.

[7]We also have data on the 2015 intra-decennial census, but inter-decennial census only surveyed selected townships and do not have crucial information on household income.

Table 1: Summary Statistics

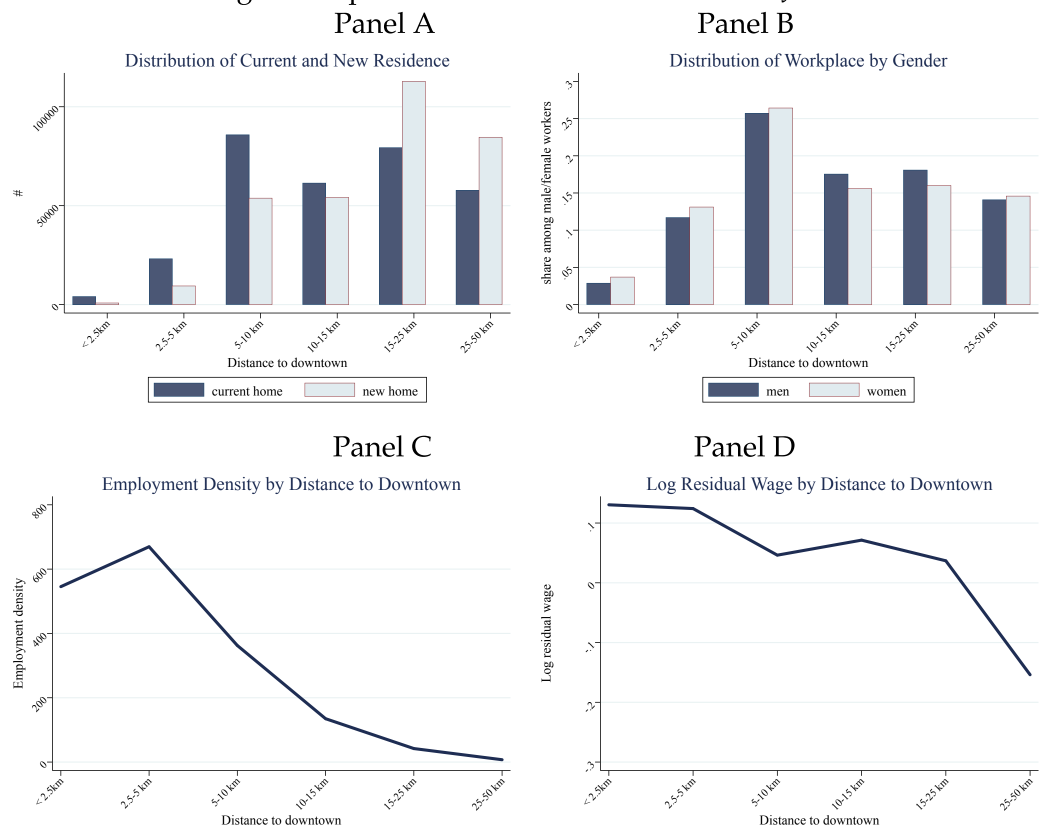
Panel A: Mortgage data				
	Households			
	mean	s.d.		
household monthly income (yuan)	9988.48	6702.61		
current home to downtown distance (km)	21.27	18.59		
new home to downtown distance (km)	23.55	16.19		
	Individuals			
	men		women	
	mean	s.d.	mean	s.d.
age	35.46	7.52	34.02	7.38
share with college degree	0.48	0.50	0.43	0.50
individual income (yuan)	5913.18	5142.10	4075.30	4020.64
work location to downtown distance (km)	17.87	17.37	17.48	17.31
current home to work distance (km)	11.79	13.96	10.48	12.66
new home to work distance (km)	15.72	14.29	14.27	13.44
Panel B: 2015 Household Travel Survey				
	men		women	
	mean	s.d.	mean	s.d.
age	44.44	9.90	41.35	8.79
share with college degree	0.32	0.47	0.36	0.48
share single	0.07	0.25	0.07	0.25
share with young children under 12	0.21	0.41	0.21	0.41
commute distance (km)	6.36	8.51	5.02	7.13
commute time (mins)	34.73	38.88	32.24	44.00
main transportation mode				
walk	0.21	0.41	0.26	0.44
bike	0.25	0.43	0.29	0.45
bus	0.12	0.32	0.16	0.37
subway	0.09	0.28	0.10	0.30
car	0.34	0.47	0.19	0.39
share by annual household income (yuan)				
<50,000	0.26	0.44	0.25	0.43
$[50,000, 100,000)$	0.45	0.50	0.45	0.50
$[100,000, 150,000)$	0.18	0.39	0.19	0.39
$[150,000, 200,000)$	0.06	0.24	0.07	0.25
$\geq 200,000$	0.04	0.20	0.05	0.21

Note: The mortgage sample includes 173,596 couples, while the 2015 Household Travel Survey covers 28,366 individuals with commuting trips, comprising 16,309 men and 12,057 women.

2.3 Basic Patterns of the Gender Commute Gap

Using the mortgage data, Table 2 reports the gender gaps in commute *distances* from the *current* homes. In Columns 1 to 4, we regress the log commute distance on a gender indicator, while controlling for various individual and household characteristics. Men's commutes are on average 8.5% longer than women's

Figure 1: Spatial Distribution of Homes and Jobs



Note: Authors' calculations using the mortgage data. In Panel D, the log wage is first residualized by regressing on a set of gender, education attainment, and age bin indicators.

(Column 1). Conditional on individual characteristics, including a set of education attainment indicators and a set of age bin indicators, the gender gap increases to 14% (Column 2). To account for the potential unobservable heterogeneity resulting from the neighborhood sorting—for example, workers who prefer to live in bigger houses may tend to live farther from downtown and have longer commutes, and male workers on average have a stronger preference for big houses than women—in Column 3 we control for a set of 18 district/county indicators for where the current home is located. Column 4 investigates the gender commute gap within the same household (comparing the husband with the wife). Controlling for household fixed effects allows us to control for the neighborhood sorting and household preferences. The gender commute gap remains at 11%.

The two remaining columns show the gender earnings gap and its relation with commuting. Female workers earn 26% less than male workers with the same qualifications. However, when it is conditional on the log commute distance, the gender pay gap declines by 9% to 23%. The magnitude of the decline in the gender pay gap associated with commuting is similar to that found in recent studies.[8] In addition, Column 6 shows that the commute distance is positively associated with the log salary, which is a com-

[8] Gutierrez (2018) finds that 10% of the gender pay gap among childless workers and more than 23% of the wage decline attributed to the child penalty can be explained by sex differences in commuting patterns. Le Barbanchon et al. (2019) estimate that around 10% of the gender wage gap is accounted for by gender differences in the willingness to pay for a shorter commute. Liu and Su (2020) find that differential commuting choices account for between 16% and 21% of the gender wage gap.

mon pattern found in many countries. Conditional on the demographic characteristics and district of the home location, a 1 log point increase in commute distance is associated with a 0.03 log point increase in monthly salary.

Table 2: Gender Gap in Commute and Earning

	(1)	(2)	(3)	(4)	(5)	(6)
	ln dist. btw. work and current home				ln salary	
= 1 if male	0.085	0.139	0.136	0.106	0.257	0.234
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
ln dist btw work and						0.032
current home						(0.001)
Fixed effects						
age bin		X	X	X	X	X
detailed edu attainment		X	X	X	X	X
current home district			X		X	X
household				X		
N of obs.	298163	298163	298163	252310	287611	280514

Note: The sample is from the mortgage data and includes workers who report a workplace and have a positive salary. The commute distance is the linear distance between the current home address and the work address. Robust standard errors are in parentheses.

3 Home Location Choice and the Gender Commute Gap

3.1 Baseline results

Table 2 does not indicate the source of the gender commute gap. The observed gender gap could be due to women only searching for jobs closer to home or systematically choosing to live closer to their places of work.

In this paper, we aim to confirm one particular explanation of the gender commute gap, namely, when deciding where to live, households tend to choose locations closer to the wife's workplace. Using the mortgage data, Table 3 shows that the commute from the *new* home is also systematically shorter for the wife. Specifically, Table 3 shows that women's commutes are between 11% and 13% shorter than men's, and the results are robust to individual demographic characteristics, a set of district or county fixed effects for where the new home is located, and the comparison within the household. Compared with the estimates in Table 2, the gender commute gap from new homes is similar in magnitude to that from current homes. These results indicate that the endogenous home location choice provides an empirically relevant explanation for the observed gender commute gap.

Table 3: Home Location Choice and Gender Gap in Commuting

	(1)	(2)	(3)	(4)
	dep var: ln dist btw work and new home			
= 1 if male	0.107	0.132	0.119	0.105
	(0.003)	(0.003)	(0.003)	(0.003)
Fixed effects				
age bin		X	X	X
detailed edu attain.		X	X	X
new home district			X	
household				X
N of obs	298163	298163	298163	252310

Note: The sample is from mortgage data. Robust standard errors are in parentheses.

This finding has implications for the gender pay gap. In contrast to explanations that focus on gender differences in the labor supply and job search behavior (e.g., Le Barbanchon et al., 2019; Liu and Su, 2020), which suggest that the gender commute gap is a *cause* of the gender pay gap, our results suggest that collective household decisions systematically favor an easy commute for the wife, and could potentially improve women's labor market participation and thus reduce the gender pay gap.

It is worth emphasizing that our finding does not rule out explanations based on the labor supply and job search behavior. In fact, both explanations stem from the fact that women derive a higher disutility from commuting. This higher disutility could be due to taste, as women's distaste for commuting might be stronger than men's, or to the household division of labor, as women's time is more tightly constrained between home and work, and a long commute may be particularly costly for them. Nevertheless, research on the gender commute gap does not explicitly consider household joint decisions, which we show play a role in the observed gender gap in commutes.

There are reasons to believe that the labor supply of women, given their disproportional obligations in home production, is likely to be more sensitive to long commutes. However, it is difficult to gauge the size of the impact on women's labor market outcomes by a household's collective choice to live closer to the wife's workplace. To answer this question in our setting, we need exogenous variation in women's commute distances and to know how much shorter commutes increase women's job retention rates and wages, which is beyond the scope of this paper. In an interesting paper that is also set in Beijing, Liu et al. (2017) show that wives in households that did not win a lottery to buy a car were less likely to be working.[9] In our setting, we show in Section 3.4 that women are not more likely to experience job turnovers than men, even though new homes are on average 40% farther away from the current workplaces.

A household's choice of living closer to the wife's workplace may well be only one factor in a household's collective decision. The husband may only agree to live closer to the wife's workplace *in exchange*

[9] To reduce traffic congestion, the municipal government of Beijing limits the number of new car licenses issued every month. The new licensees are decided by a monthly lottery. The odds of winning the lottery are low. For example, in October 2018, about 3 million individuals entered the lottery, and only 6,402 new licenses were issued.

for the wife agreeing to take on more housework. In other words, the gender commute gap may reflect the household division of labor and the intra-household bargaining. We explore these possibilities in Section 4 by looking at differences in gender commute gap across different types of households.

3.2 Commute Distances and Home Choices

Table 3 shows that from the chosen home location, the wife's commute distance is systematically shorter than the husband's, but it does not explicitly indicate *how* households choose homes. The simple analysis presented above implicitly assumes that a household can choose any geographic point in Beijing as the home location and other features of the house only affect the household's choice independent of the couple's commute distances. In reality, housing choices are discrete, and homes differ in features such as location, price, age, and amenities. These features may exhibit particular distributions in space, thus confounding the gender gap in commute distances.

In this subsection, we estimate a simple discrete choice model that explicitly focuses on the choice set households face and their housing choice decisions. The unit of choice is a "residential community." In urban China, homes are typically sold as apartment units in building complexes. Developers acquire permits from local governments to develop plots of land on which they construct several apartment buildings. These buildings form a "community." Although developers usually offer several different floor plans, buildings in the same community are typically constructed at the same time with similar designs. Homeowners in the community have access to the same amenities under the same property managers. Thus, deciding on a community is the most important step in choosing a home.[10] In our dataset, a median-sized community covers 4 hectares of land and consists of 10 buildings and a total of 900 units.

Our data comprise the universe of mortgages in Beijing form a leading lender between 2005 and 2014. As our sample covers a substantial share of the total home sales, we define the choice set each household faces as all of the communities that have a transaction record in our sample in the same year when the mortgage was obtained. There are 1,839 unique communities in 2006, the first full year in our sample, and 3,871 unique communities in 2013, the last full year in the sample. We assume that the supply of housing units in any community in the choice set is abundant, and that households are price-takers and treat community-level amenities as exogenously given.[11]

The mortgage data offer a rich set of community-level characteristics. From the exact locations of the communities and the couples' workplaces, we can calculate the commute distances for the husbands and wives from each community. These are our key variables of interest. Other community-level characteristics include the average price of one square meter of floorspace, the age of the community, and the ratio of green space. We also include the distance from each community to the *current* home to capture that households may want to live in areas they are most familiar with.

[10]Other unit-level considerations may include the floorspace, the number of stories above ground level, and the direction the unit faces. We abstract away from these considerations and assume that all units in a community are identical.

[11]This differs from the neighborhood sorting literature that treats neighborhood characteristics as endogenous (e.g., Bayer et al., 2007, Bayer et al., 2016).

We estimate the following discrete choice random utility model. A household h's utility from choosing community j is

$$u_{hj} = \beta^m \cdot \ln CommDist^m_{hj} + \beta^f \cdot \ln CommDist^f_{hj} + \rho^d \cdot \ln Price_j + X_{hj} \cdot \gamma^d + \varepsilon_{hj}, \tag{1}$$

where $\ln CommDist^m_{hj}$ and $\ln CommDist^f_{hj}$ are the log commuting distance for the husband and the wife, respectively. β^m and β^f are the parameters of primary interest. In particular, we are interested in whether the parameters are less than zero (indicating longer commutes make the community less attractive) and whether in absolute values $|\beta^f| > |\beta^m|$ (indicating that a longer commute for the wife generates a larger disutility than that for the husband). p_j is the price per square meter of community j and X_{hj} is a vector of other community characteristics, including the age of the community, the ratio of green space, and the log distance to the current home (which is household-specific). In the full model, we allow the effects of price and community-level characteristics to differ by household type d.[12] ε_{hj} is an i.i.d. shock that follows the type-I extreme value distribution.

The probability of household h choosing neighborhood j is

$$p_{hj} = \frac{\exp(u_{hj})}{\sum_k \exp(u_{hk})}. \tag{2}$$

We estimate this logit model using the maximum likelihood estimation. To reduce the computational burden, we select a subsample of 20 available communities (19 randomly chosen control communities plus the chosen community) to construct the choice set.[13]

The coefficients, $\beta = \{\beta^f, \beta^m, \rho^d, \gamma^d\}$, do not have a direct economic interpretation. Economists typically calculate the marginal effect, or the change in the probability of choosing one particular alternative by changing one unit of value in the r^{th} predictor x_{hjr}. The marginal effect can be written as

$$\frac{\partial p_{hj}}{\partial x_{hkr}} = \begin{cases} p_{hj}(1 - p_{hj})\beta_r, & \text{if } j = k \\ -p_{hj} p_{hk} \beta_r, & \text{if } j \neq k. \end{cases} \tag{3}$$

The marginal effects vary with p_{hj}, the probability of household h selecting community j, which in turn

[12] We divide the households into 12 types with three husband age groups and four household income groups. The three husband age groups are (1) under 30, (2) between 31 and 40, and (3) 41 and above. The household income groups are divided into quartiles. The household income increases rapidly during the sample period. To make the household income comparable across years, we first regress the log individual income on a set of individual characteristics, including a set of education indicators, a gender indicator, age and its quadratic term, and a set of year fixed effects. The estimated year fixed effects are then used to convert individual incomes into 2013 values. The fixed effect for 2005 is -0.6 relative to that for 2013, which means that an individual with the same characteristics, qualifications, and experience makes about 55% ($\exp(-0.6)$) less in 2005 than in 2013. House prices increase even more rapidly during the same period. To make the prices from different years comparable, we do the same type of adjustment based on observable community characteristics. The fixed effect for 2005 relative to that for 2013 is -1.364, suggesting an almost 400% increase (in nominal terms) in house price over the 8-year period. However, because the choice set consists of communities with at least one unit sold during the same year, only the variation in price across different communities in the same year matters. The results are identical with adjusted and unadjusted house prices.

[13] Specifically, for each household we keep the community that the household actually chose and randomly select 19 other communities among the communities with at least one unit sold during the same year. McFadden (1978) shows that in random utility models, estimation using choice-based sampling can asymptomatically approximate the true parameters.

depends on the values of the predictors. Because p_{hj} is bounded between 0 and 1, β_r has the same *sign* as the marginal effect. If β_r is positive, Equation 3 suggests that an increase in x_{hjr} increases the probability of option j being chosen, and because the probabilities across all options add up to 1, it reduces the probability of any other option being chosen.

In addition, the ratio of the two coefficients, $\beta_r/\beta_s,\ r \neq s$, is equal to the ratio of the two corresponding marginal effects, $(\partial p_{hj}/\partial_{hjr})/(\partial p_{hj}/\partial_{hjs})$, for *any* predictor values. This property is useful for us. For example, the magnitudes of β^f and β^m can be compared with that of ρ^d, which helps us to convert the disutility of commuting into dollar terms.

Table 4: Commute Distance and Home Choice

	(1)	(2)	(3)	(4)
	dep var: = 1 if chosen			
ln husband commute distance	-0.495	-0.560	-0.231	-0.232
	(0.004)	(0.004)	(0.005)	(0.005)
ln wife commute distance	-0.719	-0.765	-0.389	-0.393
	(0.004)	(0.004)	(0.005)	(0.005)
hhd type $\times$ community chars.				(see App. Table B.3)
ln price for 1 sqm		-0.906	-0.785	
		(0.008)	(0.001)	
ln distance to current home			-1.061	
			(0.004)	
new community (omitted category)				
community between 1-5 years old			-1.52	
			(0.011)	
community between 6-10 years old			-2.597	
			(0.012)	
community between 11-15 years old			-2.765	
			(0.013)	
community more than 16 years old			-2.912	
			(0.012)	
green ratio			0.010	
			(0.001)	
β^f/β^m	1.453	1.366	1.683	1.693
$p-$value, $H_0:\ \beta^f/\beta^m = 1$	[0.000]	[0.000]	[0.000]	[0.000]

Note: The table reports the coefficients and standard errors from the conditional fixed-effects logit model. In column 4, the coefficients associated with household type fully interact with the community characteristics are suppressed. These coefficients are presented in Appendix Table B.3.

Table 4 reports the results of estimating various versions of the conditional fixed-effects logit model presented in Equation 1. The estimated coefficients and associated standard errors are reported. Column 1 includes only the log commute distances for the husband and the wife. The commutes of both spouses have a negative sign, indicating disutility from commuting. The coefficient for the wife's log commute distance is larger in magnitude than that for the husband's. The ratio of the two coefficients, which is about 1.45 and can be rejected as statistically equal to 1, indicates the reduction in the probability of a particular community being chosen as a log point increase in the wife's commute distance is equivalent to a 1.45 log point increase in the husband's commute distance.

Column 2 adds the log unit-area price of the community, which has a negative effect on the community being chosen. Comparing the magnitudes of the coefficients suggests that the disutility of the commuting distances is sizable. A 1 log point increase in the wife's commute distance is equivalent to a 0.84 log point increase in price in terms of reducing the probability of a particular community being chosen, whereas a 1 log point increase in the husband's commute distance is equivalent to a 0.62 log point increase in price. The coefficient associated with the log commute distance for the wife remains larger than that for the husband. The ratio of the two coefficients is 1.37 and is statistically different from 1.

Column 3 adds additional community characteristics. A community is less likely to be chosen if it is farther away from the household's current home, suggesting there is a "familiarity bias." Conditional on the price, households have a strong preference for new apartments. For older communities, utility decreases with age, although the gradient is smaller as the age of the community increases. Everything else being equal, households prefer communities with larger green spaces, although it is much less important than the commute distance, price, and age of the community. The coefficients associated with the log commute distances decline by about half compared with those in Columns 1 and 2. This is largely due to the inclusion of the log distance to the current home. Households tend to live close to their places of work if they can, so the commute distance from the new home is positively correlated with the distance between the current home and new home.[14] The coefficients still suggest that the commute distance has an important influence on a household's utility from a home. A 1 log point increase in the wife's commute distance has a marginal effect on the probability of a particular home being chosen that is equivalent to about a 0.5 log point increase in price. The ratio between β^f and β^m is 1.68, which is greater than the figures in Columns 1 and 2.

Column 4 includes fully interacted terms between the community characteristics, except for the commute distances, and indicators for the 12 household types. The coefficients associated with the interaction terms are suppressed in Table 4 and instead reported in Appendix Table B.3. These coefficients are all intuitive. Price sensitivity declines with household income and age, while preferences for a new home closer to the current home are largely similar across income and age groups. Richer households also have a stronger preference for newer communities and more green space. Important for us, the coefficients associated with the log commute distances are similar to those in Column 3. This suggests that unobservable household heterogeneity is unlikely to confound our conclusion that households derive a larger disutility from the wife's commute.

[14]The correlation coefficient between the log husband's commute distance and the log distance between the current and new home is 0.67, and the correlation coefficient between the log wife's commute distance and the log distance between the current and new home is 0.7.

3.3 Robustness Checks

We now conduct robustness checks to address several concerns relating to our results. For the sake of brevity, we present the results of the specification in Table 3, Column 4. The same conclusions hold using the discrete choice models in Table 4.

3.3.1 Gender Differences in the Spatial Distribution of Jobs

The gender gap in commutes could be due to the different spatial distributions of jobs typically held by women and men. Most households in Beijing live more than five kilometers away from downtown, and the employment density and residual wage rates are highest near Beijing's city center (Figure 1, Panels C and D). If men are more likely to work in the city center while women more likely in the periphery, it is natural that women's commutes are shorter. Panel B of Figure 1 shows that this is not the case, as the spatial distributions of men's and women's jobs are similar. If anything, women are slightly more likely to work close to downtown.

To more formally test whether new homes are systematically closer to women's workplaces, we reshuffle the new homes in our sample and randomly assign them to each household. We then calculate the husband and wife's commute distances to this "fake" new home and estimate a placebo gender gap. We repeat this placebo test 1,000 times using the same specification as in Column 4 of Table 3. The coefficients associated with the male indicator are plotted in Appendix Figure B.1. The placebo gender gaps are bounded within a tight range between -0.008 and 0.003, while the gender gap using the actual home location choices is 0.105.

3.3.2 Gender Gap in Commute *Time*

Thus far, we have measured the commute length as the *linear distance* between work and home. However, commuting is costly not only because workers need to travel many miles but also because it takes time. Due to the layout of the road network and the endogenous adoption of different modes of transportation, the commute distance needs not be mapped proportionally to the commute *time*.

The mortgage data do not tell us which routes workers take, what modes of transportation they use, or how much time the commutes take. Instead, we use Baidu Maps to calculate the commute times for different modes of transportation. Specifically, we sent requests to Baidu Maps API to calculate the travel routes and times between the home and work addresses. We specify the departure time as 7am on a workday,[15] and calculate the commute time by public transportation or by car.[16]

[15]Baidu Maps uses historical road speeds together with contemporaneous variables to calculate the travel routes and times, so specifying time matters for the results. A better way may have been to specify an arrival time of 9am, when most businesses open. Unfortunately, Baidu Maps does not allow such requests. We conducted these searches in July 2020, despite our sample covering the period between 2005 and 2014. The implicit assumption is that home purchases are long-run decisions and households fully expect future changes in public transit infrastructure and road speeds. Also note that due to the COVID-19 pandemic, traffic volume in Beijing might be lower than usual. Thus, the commute speeds by car are particularly fast. Because more men drive than women, this tends to underestimate the actual gender gap in commute times in normal times.

[16]For commutes by public transit, the time walking to the bus stop or subway station is included in the total travel time from Baidu Maps. For commutes by car, we add five minutes on either end as the time needed to access to the car. For a home-work

Column 1 of Table 5 shows that the gender gap in the commute distance, measured as the length of the *route* along the road network, is about 9%. This is similar to the gender gap when the commute distance is measured as the linear distance between home and the workplace.

Columns 2 and 3 show the gender gap in commute *time* assuming workers use public transit or automobiles, respectively. The gender gap declines to 5.8% and 4%, respectively. Given the mode of transportation, the gap in time is smaller than the gap in distance because the average speed is typically higher for longer trips. In this case, the fixed cost of accessing transit and a car gets diluted in a longer trip, and longer trips are more likely to pass through suburban areas where the road speeds are higher.

One remaining issue is that the choice of transportation mode is endogenous. Although we assume that public transit is available to everyone,[17] less than 30% of Beijing households owned a car in 2019. The choice of transportation mode depends on a wide range of considerations, including car ownership, household economic conditions, access to public transport, and length of the commute. Using data from the 2015 Household Travel Survey, we estimate empirical models of the mode choice using individual and household characteristics. We then construct the same set of variables for the mortgage sample and use fitted models to predict the mode choice for each individual.[18] Men are more likely to commute by car, which is faster. Thus, the gender gap in commute time is smaller than that in commute distance. Using three different prediction models, Columns 4 to 6 of Table 5 show the gender gap in commute time is between 3.3% and 3.5% and remains statistically significant.

These results suggest a modest gender gap in commute time. The average imputed one-way commute time in the sample is 62 minutes. A 3.5% gender gap in commute time translates to 4.3 minutes less commuting time per day for women. The actual cost of an additional 4.3 minutes may be larger than it appears. The travel times for longer commutes, especially for those by car or bus, are also more likely to be unpredictable, and unpredictability imposes significant costs on commuters.

location pair with a linear distance below 1 km, we assume the worker rides a bike and impose a speed of 10 kph for the commute.

[17]Public transit in Beijing is inexpensive. A 10 km ride on a bus costs 1 *yuan* (0.14 USD), and a 6 km ride on the subway costs 3 *yuan* (0.42 USD).

[18]Appendix A describes the details of the prediction.

Table 5: Home Location Choice and Gender Gap in Commuting: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
		commute time (min)				
	road distance			predicted mode		
	(km)	transit	car	model 1	model 2	model 3
= 1 if male	0.092	0.058	0.040	0.035	0.033	0.035
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed effects						
age bin	X	X	X	X	X	X
detailed edu attain.	X	X	X	X	X	X
household	X	X	X	X	X	X
N of observations	230814	230814	230806	230806	230806	230806

Note: Robust standard errors are in parentheses. The commute time and road distance are calculated from Baidu Maps API, assuming usual traffic on a Monday morning departing at 7am. We assume that all commuters cycle to work when the linear commute distance is less than 1 km. In Columns 4 though 6, Model 1 includes individual demographic characteristics, household income, and the commute length, Model 2 adds access to public transit, and Model 3 further adds spousal characteristics. See Appendix A for details.

3.3.3 Evidence from the Household Travel Survey

We use the 2015 Beijing Household Travel Survey to provide corroborative evidence. From the survey's travel diary, we are able to observe both the commute times and mode choices, which can be directly used to estimate the gender commute gap.

The first two columns of Table 6 show that in the travel survey sample, the gender gap in commute *distance* is 20%.[19] The next two columns show that the gender gap in commute *time* is about 10%. The narrower gender gap in commute time than in commute distance is expected because men are more likely to commute long distances, and people tend to use faster modes of transportation, such as cars and the subway, for long commutes.

Another issue affecting the disutility from commuting is the amenity of the transportation mode. Intuitively, sweating in a crammed bus for an hour likely produces more disutility than sitting in an air-conditioned private car through the traffic. To account for the amenity differences across different transportation modes, we further control for mode fixed effects. Columns 5 and 6 show that the gender gap in commute time decreases to about 4%. These estimates limit the gender comparisons to the same mode of transportation. As men are more likely to use faster modes of transportation *and* have longer commutes, these estimates should be interpreted as the lower bound of the gender commute gap.

The disadvantage of the travel survey is that it is a cross-sectional dataset and does not provide the same kind of empirical setting as the mortgage data, in which we can analyze households' home location choices while holding the work locations as given. From the travel survey we do not know whether the

[19] The average commute distance in the Household Travel Survey is calculated as the linear distance between the centroids of the work TAZ and home TAZ. The travel survey over-samples those who live close to the city center. The average commute distance in the survey is 5.8 km, whereas the average commute distance in the mortgage sample is 11.1 km from current homes and 15.0 km from new homes.

gender commute gap is due to home location choices or job search behavior. The survey does record when a household moved into the current home. To make the result more comparable to the mortgage data, we restrict the sample to households that have moved in the past 5 years.[20] The last three columns of Table 6 show that the gender gap in commutes, measured by the straight-line distance, and travel time with or without being conditional on the mode of transportation, are similar to those from the larger sample of the survey and mortgage data.

Table 6: Gender Gap in Commute Time, from the Household Travel Survey

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	all households						recently moved		
	ln commute distance (km)		ln commute time (min)				ln distance (km)	ln time (min)	
= 1 if male	0.214	0.222	0.102	0.099	0.040	0.040	0.134	0.055	0.037
	(0.011)	(0.011)	(0.008)	(0.007)	(0.006)	(0.006)	(0.034)	(0.023)	(0.018)
Fixed effects									
age bin	X	X	X	X	X	X	X	X	X
detailed edu	X	X	X	X	X	X	X	X	X
home TAZ	X		X		X				
household		X		X		X	X	X	X
trans. mode					X	X			X
N of obs.	73759	72909	77698	75339	77698	75339	5803	5771	5771

Note: Each observation is a commute to or from home. Standard errors in brackets are clustered at the person level.

3.3.4 Discussion of the External Validity

One concern with our results from the mortgage data is that they are derived from a very specific population. Indeed, our sample includes those who bought homes in Beijing and took out a mortgage from a specific lender. To qualify a mortgage, the lender imposes specific qualifications. Beijing has one of the highest house price-to-income ratios in the world. In 2019, the median home price was about 40 times the average full-time annual salary. Even in the mortgage sample, among those who eventually could afford to buy a home in Beijing, the median home price is equivalent to 7 years of household income in 2005, and 9.5 years of household income in 2014. Due to the high house prices, it is reasonable to assume that households that have monthly mortgage obligations cannot afford to lose one income. To the extent that women's labor force participation decisions are more sensitive to commuting, a household may thus choose to live closer to the wife's workplace.

Although our sample does not represent Beijing's overall population, our results are robust to various subsamples with different demographics. Appendix Table B.4 shows that the gender commute gap is observed, and the magnitudes are largely comparable, across different education and age groups. It is unlikely that the overall pattern is unique to this specific sample.

[20] The results are similar with more recently-moved households, although the statistical power of the estimates declines with the sample size.

xThe systematic gender commute gap is also not unique to Beijing, which is among the largest and richest metropolises in China. The 2015 intra-decennial census is the first nationwide survey to include questions on commute times. Using a sample from this census, Appendix Figure B.2 documents that the gender gap in commute time is present in essentially all province-level administrative areas in China. In fact, Beijing has among the smallest gender gap. The estimated gender gap in commute *time* is about 3%, which is similar to the figures observed in the mortgage sample and Household Transportation Survey. At the national level, the gender gap is about 8%. These estimates are robust to the inclusion of various sets of fixed effects, including controlling for the specific transportation modes.

3.4 Job Changes after Obtaining a Mortgage

Changes of residence are often accompanied by job changes. We measure the gender gap in commute distances between new homes and the *current* jobs. It will challenge our findings if households move to new homes that are closer to their *future* jobs. In particular, the gender gap in commute distance may be the reverse of what we find if after taking out a mortgage, husbands disproportionately change their jobs to be closer to new homes.

To investigate this alternative explanation, we obtain the monthly panel of employer-employee linked records. The administrative records on employment relationships are derived from monthly compensation records. The records we have include all individuals in our mortgage sample, either as a principal borrower or as a co-borrower, provided that the individual was ever employed in the formal sector between 2006 and 2014. We also have the employer's ID for each employer-employee relationship.

We then match this monthly employer-employee panel to our main mortgage data, which contain more detailed information on the individual and firm characteristics. From the mortgage data we know the detailed addresses of the employers when the household took out a mortgage. We use this information and the employer ID to fill out the work locations in the employer-employee panel. As long as there is one employee who took out a mortgage between 2005 and 2014, we can fill out the work location for all the employees of this employer during the sample period. We are able to identify 33,816 of the 51,280 employers that show up in the monthly employer-employee panel. From those employers we are able to recover detailed addresses for about 97% of the monthly employer-employee pairs in our data. Unidentified employers are likely small firms and only account for 3% of the monthly observations.

We first examine how taking out a mortgage affects the probability of job turnovers for men and women. We define a *job switch* as a person changing employers from one month to the next, and we define a *"job loss"* as a person being in the employer-employee matched panel in one month but not in the next. Here, job loss is in quotation marks because dropping out from the panel does not necessarily mean that person became jobless. It may indicate that the person did not receive compensation for other reasons (e.g., on vacation, on sick or parental leave, etc), or the person switched to the informal sector. We restrict job losses only to cases in which a worker is not found in any employment relationship for at least three consecutive months. The results are not sensitive to the choice of the number of months to define a job loss. We group the monthly observations at the year level and generate two binary variables

Table 7: Summary Statistics of Employment Dynamics

Panel A: employment dynamics				
	men		women	
	before mortgage	after mortgage	before mortgage	after mortgage
prob. of job switches	0.158	0.112	0.152	0.100
	(0.365)	(0.315)	(0.359)	(0.300)
prob. of job losses	0.096	0.047	0.098	0.045
	(0.295)	(0.211)	(0.297)	(0.208)

Panel B: commute distance to new home (km)		
year since mortgage take-up	men	women
0	12.700	11.339
	(13.580)	(12.526)
1	12.746	11.301
	(13.576)	(12.494)
2	12.814	11.281
	(13.588)	(12.461)
3	12.886	11.311
	(13.626)	(12.484)
4	12.934	11.337
	(13.663)	(12.509)
5	12.965	11.373
	(13.651)	(12.561)

Note: Standard deviations are in parentheses. In Panel A, the sample includes all men and women that were employed during the sample period. The sample comprises 136,346 men and 124,665 women. In Panel B, the sample includes individuals for whom we observe six consecutive years of employer addresses since the year in which a mortgage was obtained. The sample comprises 33,307 men and 27,564 women.

for whether a particular worker experienced a job switch or a job loss in the year.

Panel A of Table 7 shows that the probability of job switching and that of job loss both decline after obtaining a mortgage. Before obtaining a mortgage, 15.8% of men switch jobs while about 9.6% lose their jobs in a year, whereas after obtaining the mortgage, the annual probability of job switching and job losses is 11.2% and 4.7%, respectively. Similarly, for women, the annual probability of job switching drops from 15.2% to 10.0%, and that of job loss decreases from 9.8% to 4.5%. The decline in job turnover rates after obtaining a mortgage is not necessarily causal. As workers age, the job turnover rate tends to decline (Keane and Wolpin, 1997).

More formally, we estimate the following equations:

$$y_{it} = \sum_{\tau \neq -1} \beta_\tau \cdot \text{Yr\_to\_Mort}_{it}^{\tau} + \lambda_i + \theta_t + \varepsilon_{it}, \tag{4}$$

$$y_{it} = \sum_{\tau \neq -1} \beta_\tau \cdot \text{Yr\_to\_Mort}_{it}^{\tau} + \sum_{\tau \neq -1} \gamma_\tau \cdot \text{Yr\_to\_Mort}_{it}^{\tau} \cdot \text{male}_i \qquad (5)$$
$$+ \rho \cdot t \cdot \text{male}_i + \lambda_i + \theta_t + \varepsilon_{it}.$$

In Equation 4, y_{it} is a binary variable indicating whether person i experienced a job switch or a job loss in year t. $\text{Yr\_to\_Mort}_{it}^{\tau}$ is a binary variable indicating year τ since the household obtained a mortgage (τ can be negative). The coefficient associated with the year before the mortgage was obtained is normalized to 0. λ_i is the person fixed effect. θ_t is the year fixed effect. Equation 5 further investigates the gender difference in the employment dynamics. male_i is an indicator for men. $t \cdot \text{male}_i$ captures the linear trend in the employment dynamics that could differ by gender. The main parameters of interest is γ_τ.

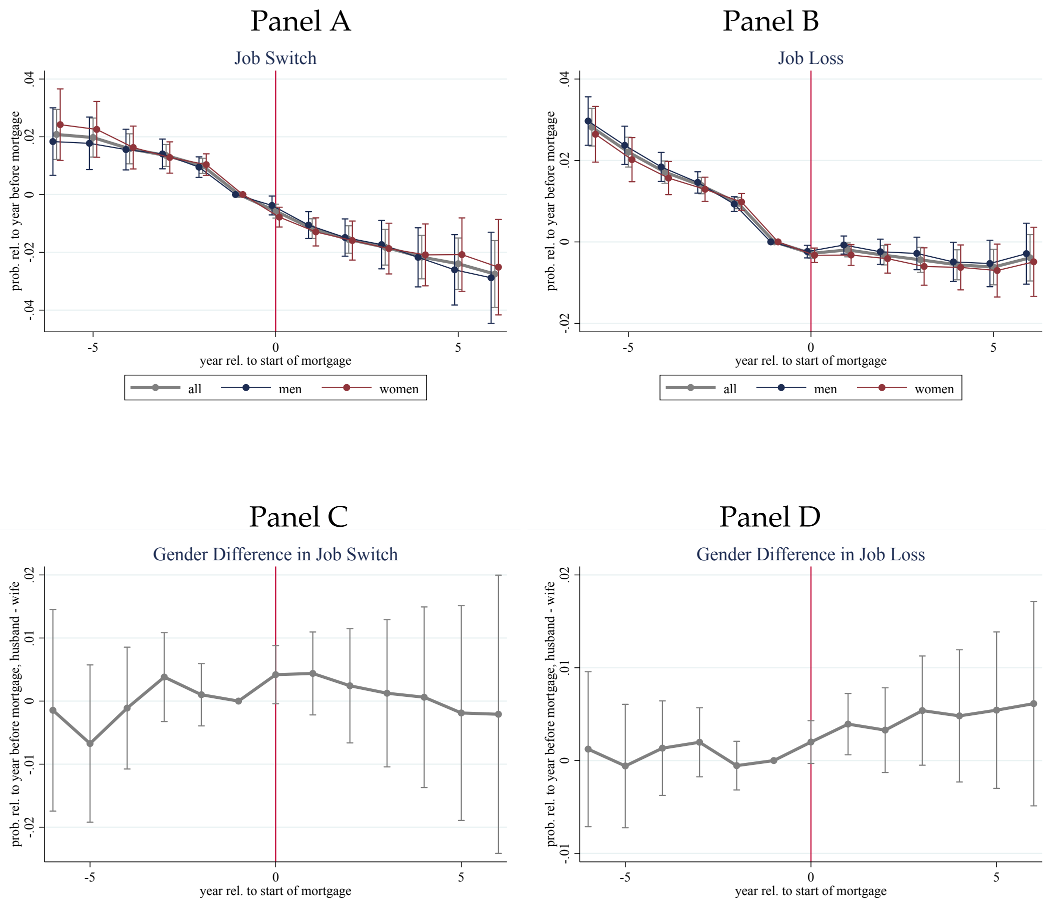
The first two panels of Figure 2 report the results of Equation 4 where the dependent variable is job switch (Panel A) and job loss (Panel B). For both men and women, the job switch and job loss probabilities both decrease, which is natural as workers age. There is no obvious trend break for job switches around the time of a mortgage is obtained, while the probability of job loss seems to level off after a mortgage is taken out. This suggests a "job-locking" effect of mortgage – when serving a mortgage, losing a job is more consequential and even job switches are more risky. Important for us, Panels C and D report the estimates of γ_τ in Equation 5 and show that there are no discernible differences between men and women in job turnover rates both before and after taking out a mortgage. If anything, women's job loss rates appear to decline relative to men's after obtaining a mortgage.

We then investigate how the gender gap in commute distance evolves after a household takes out a mortgage. Panel B of Table 7 uses a fully-balanced panel of individuals for whom we observe six consecutive years of employer addresses since taking out a mortgage. The results show that men's average commute distances tend to slightly *increase* after taking out a mortgage. By the 5$^{\text{th}}$ year, men's average commute distance is 2% longer than that when the mortgage was taken out. For women, the commute distance is largely unchanged.

More formally, we estimate versions of Equations 4 and 5 but restrict the sample period to the post-mortgage take-out year (the year the mortgage is obtained is the leave-out year) and replace the dependent variable with the log commute distance between the new home and the year-end work location. Panel A of Figure 3 shows that relative to the year in which a mortgage is taken out, people gradually work in locations that are *farther away* from their new homes. By the end of the third year since obtaining a mortgage, the commute distances are approximately 1% longer, with men's commute distances increasing more than women's. Panel B shows that compared with the gender gap in commute distance in the year a mortgage is taken out, the gender gap increases by 1% by the third year and by 1.3% in year five. These effects are precisely estimated, but small in magnitude. Overall, the evidence presented in this subsection shows that our main results on the gender commute gap remain consistent, and even increase over the next few years after a mortgage is taken out.[21]

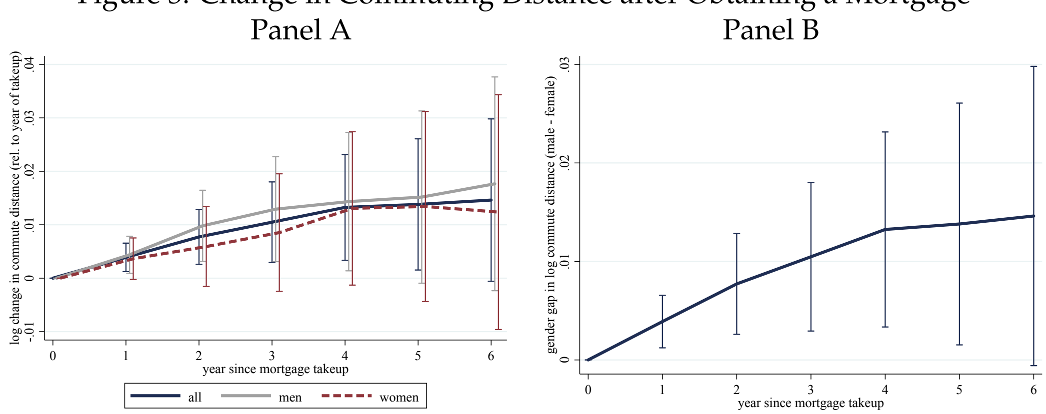
[21] An interesting and related exercise would be to investigate whether moving to a new home farther away from the workplace increases job turnover, and whether the effect differs by gender. We leave this for future work because exogenous variations in the commute distance from the new home are needed for this exercise. In this paper, the commute distance from the new home is treated as endogenous.

Figure 2: Employment Dynamics around Obtaining a Mortgage



Note: Panels A and B are from estimating Equation 4. Panels C and D are from estimating Equation 5. The vertical bars indicate 95% confidence intervals.

Figure 3: Change in Commuting Distance after Obtaining a Mortgage



Note: The vertical bars indicate 95% confidence intervals.

4 Potential Mechanisms

4.1 A Simple Model of Home-Location Choice

To interpret the empirical findings presented in the previous section, we now build an intra-household bargaining model to illustrate how a couple jointly determines their home location. Consistent with our empirical setting, the model takes the work locations of the husband and wife as given. We show that the commuting times for both the husband and wife play a crucial role in their home location choice and the intra-household allocation of time and goods.

A household h consists of a husband (m) and a wife (f). The couple jointly decide where to live. We assume that the work locations of the husband and wife are fixed, and therefore, the home location determines the commute time for the husband t_m and for the wife t_f. For simplicity, we assume that the couple choose a home located on the straight line between the two workplaces, and hence, the sum of the commuting times of the husband and wife (t_m and t_f) is a fixed number D.[22]

In addition, each member of the couple chooses their own private good c_i, home production time l_i, and leisure L_i. They derive utility from their own private good (c_i), a household public good (Q), and leisure (L_i). Production of the household public good requires labor input from the husband (l_m) and the wife (l_f).

The utility function of the household is

$$
\begin{aligned}
&U_m + \mu U_f, \\
&\text{where } U_i = \log C_i + \alpha_Q \log Q + \alpha_i \log L_i \quad \text{for } i \in \{m, f\}
\end{aligned}
\tag{6}
$$

μ is the Pareto weight of the wife. α_Q captures the preference on public good Q. α_i capture individual i's idiosyncratic preference for leisure L_i.

Household public good Q is produced by husband's time l_m and wife's time l_f:

$$Q = (\gamma l_f^{\rho} + l_m^{\rho})^{1/\rho}. \tag{7}$$

Individual time inputs are aggregated to household public good via a constant-elasticity-of-substitution (CES) production function. γ captures the time efficiency of wife relative to that of husband. ρ describes the elasticity of substitution between wife's time and husband's time.

We assume that each worker supplies a fixed amount of time in working.[23] The time constraint for individual i is:

$$T = t_i + l_i + L_i. \tag{8}$$

T is the total time net of working hours. T is divided into commute time t_i, time spent on producing the

[22] We assume that residential amenities are the same anywhere in the city. With costly commuting being the only consideration, the household first minimizes the total commuting distance. Any point outside of the straight line between the two workplaces is suboptimal.

[23] We argue that this is a reasonable assumption for those in the mortgage sample, who typically need to be full-time workers to qualify for a mortgage.

household public good l_i, and leisure L_i.

Household budget constraint is:

$$w_m + w_f = c_m + c_f + bt_m + bt_f, \tag{9}$$

where w_i is the labor earnings of spouse i, and b is the monetary cost of commuting.

The household's problem is as follows. Conditional on the work locations of the couple, the household solves the resource allocation problem by choosing the commuting time, home production time, leisure time, and consumption level for the husband and the wife (t_m, t_f, l_m, l_f, L_m, L_f, c_m, and c_f):

$$\begin{aligned}
\max_{t_m,t_f,l_m,l_f,L_m,L_f,c_m,c_f} \quad & \log c_m + \mu \log c_f + (1+\mu)\alpha_Q \log Q + \alpha_m \log L_m + \mu\alpha_f \log L_f \qquad (10)\\
s.t. \quad & t_i + l_i + L_i = T \quad \text{for, } i \in \{m, f\}\\
& w_m + w_f = c_m + c_f + bD\\
& Q = (\gamma l_f^{\rho} + l_m^{\rho})^{1/\rho}\\
& t_m + t_f = D\\
& t_i \geq 0, \quad l_i \geq 0, \quad L_i \geq 0 \quad \text{for } i \in \{m, f\}
\end{aligned}$$

The first order conditions with respect to c_m and c_f are:

$$\begin{aligned}
\frac{1}{c_m} &= \lambda_c\\
\mu\frac{1}{c_f} &= \lambda_c\\
\frac{c_m}{c_f} &= \frac{1}{\mu} \qquad (11)
\end{aligned}$$

Given household's income $w_m + w_f$, the couple splits the consumption based on Equation 11. Therefore, we can solve c_m and c_f:

$$\begin{aligned}
c_m &= \frac{1}{1+\mu}(w_m + w_f - bD)\\
c_f &= \frac{\mu}{1+\mu}(w_m + w_f - bD)
\end{aligned}$$

Plugging the solution back into the household maximization problem, we have the following Lagrange equation:

$$\max_{h_m,h_f,t_m,l_m,l_f,L_m,L_f} \log \frac{1}{1+\mu}(w_m + w_f - bD) + \mu \log \frac{\mu}{1+\mu}(w_m + w_f - bD) \tag{12}$$
$$+ (1+\mu)\alpha_Q \rho \log(\gamma l_f^\rho + l_m^\rho) + \alpha_m \log L_m + \mu \alpha_f \log L_f$$
$$+ \lambda_m (T - l_m - L_m - t_m) + \lambda_f (T - l_m - L_m - D + t_m)$$

Here we replace t_f with $D - t_m$ and replace L_i with $T - l_i - t_i$.

Taking the first order condition with respect to l_m, l_f and t_m, we have

$$l_m: \quad \frac{(1+\mu)\alpha_Q l_m^{\rho-1}}{\gamma l_f^\rho + l_m^\rho} = \lambda_m \tag{13}$$

$$l_f: \quad \frac{(1+\mu)\alpha_Q \gamma l_f^{\rho-1}}{\gamma l_f^\rho + l_m^\rho} = \lambda_f \tag{14}$$

$$L_m: \quad \frac{\alpha_m}{L_m} = \lambda_m \tag{15}$$

$$L_f: \quad \mu \frac{\alpha_f}{L_f} = \lambda_f \tag{16}$$

$$t_m: \quad \lambda_m = \lambda_f \tag{17}$$

From Equations 13 and 14, we have

$$\frac{l_m}{l_f} = (\gamma)^{\frac{1}{\rho-1}} \tag{18}$$

From Equations 15 and 16, we have

$$\frac{L_m}{L_f} = \frac{\alpha_m}{\mu \alpha_f} \tag{19}$$

Plugging Equations 18 and 19 into the time constraint, we get

$$l_m + l_f + L_m + L_f = 2T - D$$
$$(\gamma^{\frac{1}{1-\rho}} + 1) l_m + (1 + \frac{\mu \alpha_f}{\alpha_m}) L_m = 2T - D \tag{20}$$

Plugging Equations 15 and 18 into Equation 13, we get

$$\frac{(1+\mu)\alpha_Q}{(\gamma^{\frac{2-\rho}{1-\rho}} + 1) l_m} = \frac{\alpha_m}{L_m}$$
$$\alpha_m (\gamma^{\frac{2-\rho}{1-\rho}} + 1) l_m - (1+\mu)\alpha_Q L_m = 0 \tag{21}$$

Solving Equations 20 and 21, we get the solution of l_m and L_m:

$$l_m = \frac{(2T-D)(1+\mu)\alpha_Q}{(\gamma^{\frac{1}{1-\rho}}+1)(1+\mu)\alpha_Q + \alpha_m(\gamma^{\frac{2-\rho}{1-\rho}}+1)(1+\frac{\mu\alpha_f}{\alpha_m})} \tag{22}$$

$$L_m = \frac{(2T-D)\alpha_m(\gamma^{\frac{2-\rho}{1-\rho}}+1)}{(\gamma^{\frac{1}{1-\rho}}+1)(1+\mu)\alpha_Q + \alpha_m(\gamma^{\frac{2-\rho}{1-\rho}}+1)(1+\frac{\mu\alpha_f}{\alpha_m})} \tag{23}$$

From the above two equations and the time constraint, we get

$$t_m = T - \frac{(2T-D)((1+\mu)\alpha_Q + \alpha_m(\gamma^{\frac{2-\rho}{1-\rho}}+1))}{(\gamma^{\frac{1}{1-\rho}}+1)(1+\mu)\alpha_Q + \alpha_m(\gamma^{\frac{2-\rho}{1-\rho}}+1)(1+\frac{\mu\alpha_f}{\alpha_m})} \tag{24}$$

The model generates the following predictions regarding how the husband's commute time changes with the wife's bargaining power, wife's home production efficiency, and the value of the couple's public goods. Note that the model predicts that a change in the earnings of the husband and wife has no direct effect on the commute times. The relative earnings of the husband and wife can affect the couple's bargaining power, which in turn can affect the spouses' commute times.

Proposition 1 *When the value of public goods α_Q increases, husband's commuting time t_m increases if $\gamma^{\frac{1}{1-\rho}} > \frac{\mu\alpha_f}{\alpha_m}$.*

Proof: Based on Equation 24, t_m increases with α_Q if and only if:

$$\frac{(1+\mu)}{(\gamma^{\frac{1}{1-\rho}}+1)(1+\mu)} < \frac{\alpha_m(\gamma^{\frac{2-\rho}{1-\rho}}+1)}{\alpha_m(\gamma^{\frac{2-\rho}{1-\rho}}+1)(1+\frac{\mu\alpha_f}{\alpha_m})}$$

This is equivalent to

$$\gamma^{\frac{1}{1-\rho}} > \frac{\mu\alpha_f}{\alpha_m}$$

This condition is likely to hold because $\gamma > 1$ (women are more efficient in home production than men) and $\rho < 1$, so $\gamma^{\frac{1}{1-\rho}} > 1$. As long as μ and α_f/α_m are not much bigger than one, the condition holds.

Proposition 2 *When wife's home production efficiency γ increases, husband's commuting time t_m increases if γ is no greater than a threshold γ^*.*

Proof: Take the derivative of t_m with respect to γ (based on Equation 24), we have

$$\frac{\partial t_m}{\partial \gamma} = -\frac{(1+\mu)\alpha_Q[(1-\rho)\alpha_m\gamma^{\frac{2-\rho}{1-\rho}} - (2-\rho)\mu\alpha_f\gamma - ((1+\mu)\alpha_Q+\alpha_m)]}{((\gamma^{\frac{1}{1-\rho}}+1)(1+\mu)\alpha_Q + \alpha_m(\gamma^{\frac{2-\rho}{1-\rho}}+1)(1+\frac{\mu\alpha_f}{\alpha_m}))^2}$$

Given that $(1-\rho)\alpha_m\gamma^{\frac{2-\rho}{1-\rho}}$ is an exponential function increasing in γ and $(2-\rho)\mu\alpha_f\gamma + ((1+\mu)\alpha_Q + \alpha_m)$ is a linear function increasing in γ, there will be a unique turning point γ^*.[24] When $\gamma \leq \gamma^*$, $\frac{\partial t_m}{\partial \gamma} > 0$ and t_m increases with γ; otherwise, t_m decreases with γ. We examine $\frac{\partial t_m}{\partial \gamma}$ at $\gamma = 1$ and find that the derivative is positive at $\gamma = 1$, suggesting that the positive effect of μ on t_m is likely to hold in a reasonable range of $\mu \in [1, t^*]$.

Proposition 3 *When the wife's bargaining power μ increases, the husband's commuting time t_m increases.*

Proof: We can re-arrange Equation 24 as a function of μ.

$$t_m = T - \frac{\alpha_Q\mu + (\alpha_Q + (\gamma^{\frac{2-\rho}{1-\rho}} + 1)\alpha_m)}{((\gamma^{\frac{1}{1-\rho}} + 1)\alpha_Q + (\gamma^{\frac{2-\rho}{1-\rho}} + 1)\alpha_f)\mu + ((\gamma^{\frac{1}{1-\rho}} + 1)\alpha_Q + (\gamma^{\frac{2-\rho}{1-\rho}} + 1)\alpha_m)}$$

t_m increases with μ if and only if

$$\frac{\alpha_Q}{(\gamma^{\frac{1}{1-\rho}} + 1)\alpha_Q + (\gamma^{\frac{2-\rho}{1-\rho}} + 1)\alpha_f} < \frac{\alpha_Q + (\gamma^{\frac{2-\rho}{1-\rho}} + 1)\alpha_m}{(\gamma^{\frac{1}{1-\rho}} + 1)\alpha_Q + (\gamma^{\frac{2-\rho}{1-\rho}} + 1)\alpha_m}$$

This is equivalent to the following

$$\gamma^{\frac{1}{1-\rho}}(\gamma^{\frac{2-\rho}{1-\rho}} + 1)\alpha_m\alpha_Q + (\gamma^{\frac{2-\rho}{1-\rho}} + 1)\alpha_f\alpha_Q + (\gamma^{\frac{2-\rho}{1-\rho}} + 1)^2\alpha_m\alpha_f > 0$$

Because the above inequality always holds, we prove the proposition.

4.2 Empirical Evidence

4.2.1 Intra-household Division of Labor

Proposition 1 predicts that households that have a higher value of household public goods are likely to see a larger gender commute gap. Proposition 2 predicts that the gender commute gap also increases with women's home production efficiency. Households with young children are likely to have a larger demand for household products — taking care of the children. In addition, wife's home production efficiency may increase after having young children as the childcare accounts for an increasingly large share of household production, in which the mother typically is assumed to be more efficient. Based on the intra-household division of labor, the model predicts that households with young children may have a larger gender commute gap than those without.

[24] We can find γ^* by solving $(1-\rho)\alpha_m\gamma^{\frac{2-\rho}{1-\rho}} = (2-\rho)\mu\alpha_f\gamma + ((1+\mu)\alpha_Q + \alpha_m)$.

Table 8: Household Division of Labor and the Gender Gap in Commute

	(1)	(2)	(3)	(4)	(5)	(6)
	mortgage		household travel survey			
	ln commute dist.		ln commute time			
= 1 if male	0.106	0.082	0.094	0.061	0.125	0.042
	(0.004)	(0.005)	(0.008)	(0.012)	(0.009)	(0.035)
male × wife btw 30-39		0.050				
		(0.007)				
male × has children				0.056		
				(0.016)		
Fixed effects						
age bin	X	X	X	X	X	X
detailed edu	X	X	X	X	X	X
home district/TAZ					X	X
household	X	X	X	X		
subsample					married	single
N of obs.	252310	252310	59894	59894	55890	3980

Note: The regressions use the mortgage sample in Columns 1 and 2, and the sample from the Household Travel Survey in Columns 3 through 6. Robust standard errors are used in Columns 1 and 2, and standard errors are clustered at the person level in Columns 3 through 6.

The mortgage data do not include information on other family members. We focus on households with wives aged 30 and 39, as this age group is most likely to have young children at home. Columns 1 and 2 of Table 8 show that while the average gender gap in commute distances across all sample households is about 11%, the gender commute gap for households with wives aged 30 and 39 is 5 percentage points (p.p.) larger than that of the other groups.

The Household Travel Survey includes a roster of household members. Columns 3 and 4 report that the sample average of the gender gap in commute time is about 9.4%, while the gender gap for households with children at home is 5.6 p.p. higher than other households in the sample.

Another way to gauge the association between household production and the gender commute gap is to compare single workers and married couples. Presumably married couples get higher values from household production, and the gender gap in commutes is larger. Columns 5 and 6 confirm this hypothesis. The gender commute gap is 12.5% among married workers and only 4.2% and not statistically significant among single workers.

Table 9: Intra-household Bargaining and the Gender Gap in Commutes

(1)		(2)		(3)	
education		individual income		age	
$\{husband, wife\}$		$ln(husband) - ln(wife)$		$husband - wife$	
$\{H, H\}$	0.098	1^{st} quartile	0.129	< 0	0.098
	(0.004)		(0.005)		(0.006)
$\{H, L\}$	0.047	2^{nd} quartile	0.093	$[0, 1]$	0.100
	(0.007)		(0.005)		(0.004)
$\{L, H\}$	0.153	3^{rd} quartile	0.102	$[2, 3]$	0.110
	(0.008)		(0.006)		(0.005)
$\{L, L\}$	0.103	4^{th} quartile	0.070	≥ 4	0.079
	(0.004)		(0.005)		(0.006)

Note: The sample is from the mortgage data. Each observation is a household with both the husband and wife working. There are 134,613 households in the sample. The dependent variable is the log difference between the husband's commute distance and the wife's commute distance. In each column, the regressors are indicators for different types of households. The constant is suppressed. In Column 1, H and L indicate college educated and non-college educated, respectively. $\{H, H\}$, $\{H, L\}$, $\{L, H\}$, and $\{L, L\}$ account for 34%, 14%, 9%, and 43% of all households, respectively. In Column 2, the quartile cutoffs for the log differences between the husband and wife's incomes are, respectively, -0.16, 0.22, and 0.76. In Column 3, the husband is younger than the wife in about 18% of the households, while the wife is between 0 and 1 year younger than the husband in 40% of the households, between 2 and 3 years younger in 25% of the households, and at least 4 years younger in 17% of the households. Robust standard errors are in brackets.

4.2.2 Intra-household Bargaining

Proposition 3 predicts that the gender commute gap will be higher in households in which the wife has more bargaining power. We proxy intra-household bargaining power as the couple's relative education, income,[25] and age. We regress the difference in the log commute distance between the husband and wife on a set of binary variables indicating the different types of households along these dimensions.

Column 1 of Table 9 reports the gender commute gaps in households in which the husband and wife have different levels of education. The figures show that the gender commute gap is largest among households in which the wife has a college degree while the husband does not. In those households, the gender commute gap is 15%, compared with only 5% among households in which the husband has a college degree while the wife does not. For households in which the couple's education levels are the same, i.e., both or neither have a college degree, the gender commute gap is similar at 10%. This is reassuring because it indicates that the gender gap is more likely determined by the *relative* bargaining power between the husband and wife instead of their education levels.

Column 2 shows that the intra-household gender commute gap is smaller in households in which the husband-to-wife income gap is larger. Column 3 shows that the commute gap slightly increases as the husband-to-wife age gap increases, while the commute gap is smaller when the husband is much older (more than 4 years older) than the wife. These results are largely consistent with Proposition 1. However, it is worth noting that even in households in which the husband is likely to have the highest

[25] Note that the model predicts that individual income levels will not directly affect the commute gap, although they affect intra-household bargaining power.

bargaining power, the husbands still have much longer commutes than the wives.

Appendix Table B.5 reports the results for the household division of labor and within-household bargaining from the counterpart discrete choice model described in Equation 1. Across all household types, longer commute distances are associated with a smaller probability of a particular home being chosen. The coefficients associated with the wife's log commute distance are always significantly higher in absolute terms than those associated with the husband's log commute distance. Moreover, the difference is greater among families with young children, households in which the wife is better educated than the husband, households in which the wife has a higher income than the husband, and households in which the wife is younger than the husband.

5 Conclusions

Leveraging a unique dataset of home mortgages in Beijing, this paper studies whether a household's choice of home location contributes to the gender gap in commutes. We find that households' new homes are on average 10% closer to the wife's workplace than to the husband's. We find that the gender gap is larger among married workers, households with children, and households in which the wife likely has greater bargaining power. We show that these empirical results are consistent with a model of intra-household bargaining and division of labor.

Given women's traditional roles in the home, our findings indicate that household collective home location choices tend to reduce women's commutes, and thus potentially mitigate the gender gap in labor market outcomes. We find evidence suggesting that the women in our sample are not more likely to experience a job turnover than men after buying a new home, although moving to a new home on average increases the commute distance by 40%.

Our finding differs from the view that the gender gap in commutes is caused by the different job search behaviors of men and women. According to this view, women cannot or are unwilling to commute long distances and thus search for jobs closer to home, which in turn limits their labor market outcomes.

However, we argue that these two causes of the observed gender commute gap can coexist. Due to data limitations, we were unable to examine the important question of how much the household home location choice, which makes it relatively easier for the wife to commute, mitigates the gender pay gap. To answer this question would require exogenous shocks to workers' commuting times and observing subsequent changes in wages and job retention rates. To decompose the different channels that contribute to the gender commute gap, future empirical studies should exploit natural experiments in both home location choices and job searches.

References

Altonji, Joseph G and Rebecca M Blank, "Race and gender in the labor market," *Handbook of labor economics*, 1999, *3*, 3143–3259.

Barbanchon, Thomas Le, Roland Rathelot, and Alexandra Roulet, "Gender Differences in Job Search: Trading off Commute against Wage," *Available at SSRN 3467750*, 2019.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 2007, *115* (4), 588–638.

_ , Robert McMillan, Alvin Murphy, and Christopher Timmins, "A Dynamic Model of Demand for Houses and Neighborhoods," *Econometrica*, 2016, *84* (3), 893–942.

Becker, Gary, *A Treatise on the Family*, Harvard university press, 2009.

Bertrand, Marianne, Claudia Goldin, and Lawrence F Katz, "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors," *American Economic Journal: Applied Economics*, 2010, *2* (3), 228–55.

Black, Dan A, Natalia Kolesnikova, and Lowell J Taylor, "Why do So Few Women Work in New York (and So Many in Minneapolis)? Labor Supply of Married Women across US Cities," *Journal of Urban Economics*, 2014, *79*, 59–71.

Chiappori, Pierre-André, "Rational Household Labor Supply," *Econometrica*, 1988, pp. 63–90.

_ , "Collective Labor Supply and Welfare," *Journal of Political Economy*, 1992, *100* (3), 437–467.

Cortés, Patricia and Jessica Pan, "When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap," *Journal of Labor Economics*, 2019, *37* (2), 351–398.

Couture, Victor, Gilles Duranton, and Matthew A Turner, "Speed," *Review of Economics and Statistics*, 2018, *100* (4), 725–739.

Crane, Randall, "Is There a Quiet Revolution in Women's Travel? Revisiting the Gender Gap in Commuting," *Journal of the American Planning Association*, 2007, *73* (3), 298–316.

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

Gutierrez, Federico, "Commuting Patterns, the Spatial Distribution of Jobs and the Gender Pay Gap in the US," *Available at SSRN 3290650*, 2018.

Keane, Michael P and Kenneth I Wolpin, "The career decisions of young men," *Journal of political Economy*, 1997, *105* (3), 473–522.

Kuminoff, Nicolai V., Chris Parmeter, Dan Phaneuf, Jaren Pope, Chris Timmins, and Especially V, "Partial Identification of Preferences in a Dual-Market Sorting Equilibrium," 2012.

Li, Shanjun, Panle Jia Barwick, Andrew Waxman, and Jing Wu, "The Economic and Health Impacts of Subway Construction: Evidence from Beijing," 2018.

Liu, Antung A, Joshua Linn, Ping Qin, and Jun Yang, "The Effect of Vehicle Restrictions on Female Labor Supply," *RFF Report*, 2017.

Liu, Sitian and Yichen Su, "The Geography of Jobs and the Gender Wage Gap," 2020.

MacDonald, Heather I, "Women's Employment and Commuting: Explaining the Links," *Journal of Planning Literature*, 1999, *13* (3), 267–283.

Madden, Janice F., "Why Women Work Closer to Home," *Urban studies*, 1981, *18* (2), 181–194.

Marianne, Bertrand, "New perspectives on gender," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1543–1590.

McFadden, Daniel, "Modeling the choice of residential location," *Transportation Research Record*, 1978, (673).

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg, "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 2018, *108* (12), 3855–90.

Organization for Economic Cooperation and Development, "OECD family database," 2015.

Severen, Christopher, "Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification," 2019.

Tsivanidis, Nick, "The Aggregate and Distributional Effects of Urban Transit Infrastructure: Evidence from Bogotá's Transmilenio," 2018.

White, Michelle J, "A model of Residential Location Choice and Commuting by Men and Women Workers," *Journal of Regional Science*, 1977, *17* (1), 41–52.

_ , "Sex Differences in Urban Commuting Patterns," *American Economic Review*, 1986, *76* (2), 368–372.

Online Appendix for
Home Location Choices and the Gender Commute Gap
Not for Publication

A Details of Calculating Commute Time using Baidu Maps

Baidu Maps provides an developer interface at https://lbsyun.baidu.com/. For each individual in the mortgage sample, we send queries to the server to calculate commute time from the current and new home addresses to the work address. We calculate commute time using two kinds of mode choices: public transit or private car. Public transit includes buses and subways. Queries can be made based on current traffic condition or predicted traffic condition in the next week. We use the predicted traffic condition in the next Monday morning, assuming departure time at 7AM. We sent the queries in the first week of July, 2020.

Baidu Maps typically returns a few routes, avoidance, and interchange options. For trips with the public transit, the options include the one with the shortest time, shortest walking distance, fewest interchanges, or avoiding the subway (which is more expensive than buses). For automobile trips, the options include the one with the shortest time, the shortest distance, or avoiding highways. For each mode of transportation, we use the default option, which is typically the one with the shortest travel time. From trips with automobiles, commute distance through the road network is also recorded. Figure A.1 shows an example of route planning on Baidu Maps using the home and work location of an individual in the sample.

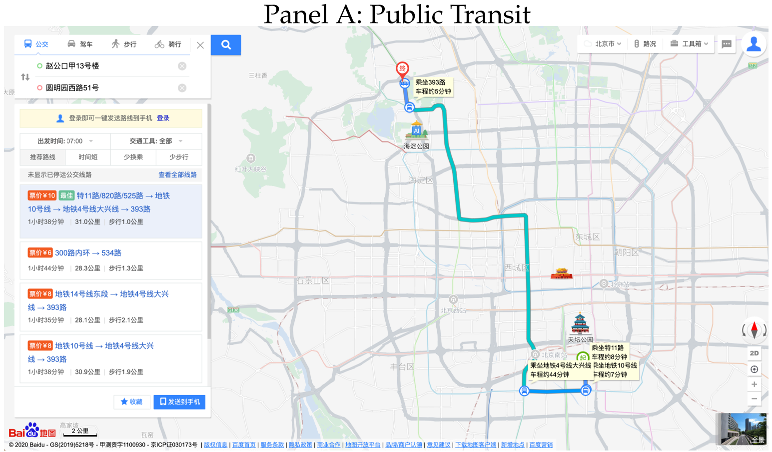
From new homes in the mortgage data, the average morning commute by private car costs 49.1 minutes (standard deviation 28.4), the average driving distance is 21.4 km. This implies an average driving speed of 23.3 kilometers per hour (km/h). The average commute time by public transit is 73.8 minutes. Assuming the road distance is the same for public transit, this results in an average speed of public transit of 15.4 km/h.

It is worth noting that by July 2020, COVID was still a concern in Beijing and demand for travel was still low. So the rush hour speed likely is much higher than in the normal days, especially for commutes by car. The 2015 Transportation Survey reveals a much slower speed. The average speed by car for commute trips in that sample was about 13 km/h (including time accessing to cars), that for commutes by bus was about 11 km/h. Because men are more likely to drive. The faster-than-usual drive speed during the pandemic leads to an underestimation of gender commute gap.

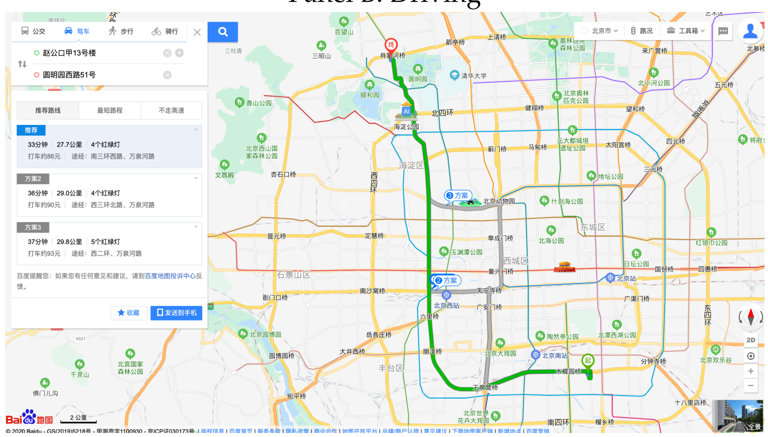
Commute time by public transit includes walking time to the bus stop or subway station, as well as expected waiting time. Commute time by private car is driving time only. We add 5 minutes on each end of a car trip to account for the access time to one's car. For commutes with linear distances below 1 kilometer, we also calculate the time needed if one chooses to ride the bike to work, assuming the bike route is the same as the route for driving and imposing an average speed of cycling in the city as 10 kilometers per hour. We assume there is no access time to one's bike. And if bike turns out to be the most time-saving option, we assume the person chooses bike.

Figure A.1: Example of Route Planning on Baidu Maps

Panel A: Public Transit



Panel B: Driving



Note: Example from Baidu Maps. https://maps.baidu.com/. Accessed on July 15th, 2020.

The mortgage data do not tell us which mode of transportation one uses for commute. We impute each individual's mode choice by fitting an discrete choice model using the Household Travel Survey, from which we observe respondents' mode choices. Specifically, we estimate a logit model where the outcome variable is an indicator for whether the commute is by car. Explanatory variables including individual's demographics, household income, commuting distance, ease of access to subway from home and workplace, and spousal characteristics. We interact each characteristic with binary gender indicators.

Table A.1 reports the marginal effects at the sample mean from three versions of logit estimations. Age has an inversed-U-shaped effect on the probability of driving. For both men and women, the age group between 35 and 44 are most likely to drive. This is probably due to the low car ownership among the younger groups, and fewer older people know how to drive, since private cars have only recently become popular among Chinese households. People with a college degree are more likely to drive even conditional on household income. Having a college degree has a particularly high marginal effect on driving for women. Not surprisingly, couples in households with higher income are more likely to drive. Husbands in households with an annual income of above 300 thousand *yuan* are 10 percentage points (p.p.) more likely to drive than those in households with an annual income less than 50 thousand *yuan*, while wives are about 20 p.p. more likely. Individuals who live farther away from their workplaces are more likely to drive. If commute distance is more than 15 kilometers, men are 35 p.p. more likely and women are 41 p.p. more likely to drive compared with commute distance that is below 1 kilometer, in which case most choose to walk or bike.

Model 2 adds measures of access to the subway from home and workplace (from centroids of the corresponding TAZ). Being close (less than 0.5 kilometers) to a subway station is strongly associated with less car use. Model 3 further adds spousal characteristics. Having a college-educated spouse is positively associated with one's own probability of using car for commute. If the spouse's commute is long, one's own probability of commuting by car is low. This is probably because most households in Beijing have at most one car.

We construct the same set of individual and household characteristics in the mortgage sample. Using these estimated marginal effects, we predict the probabilities of driving for commute for each individual in the mortgage sample. In the Household Travel Survey, 34% of men commute by car, compared with 19% of women. According to the first model, the average predicted probability for men in the mortgage sample is 50%, for women it is 33%. The second model predicts a probability of 61% for men and 51% for women. The third model predicts 60% and 57%, respectively.

For columns with predicted transportation mode in Table 5, we use the predicted commute time based on the probability of using either mode. The predicted commute time is the sum of the multiplication between the predicted probability of choosing car or public transit and the commute time of using either mode.

Table A.1: Transportation Mode Prediction Models

	model 1		model 2		model 3	
dep var: = 1 commute by car	male	female	male	female	male	female
Age						
less than 34	0.061	-0.146	0.120	-0.110	0.215	-0.060
	(0.039)	(0.024)	(0.044)	(0.028)	(0.070)	(0.052)
35 - 44	0.124	-0.119	0.180	-0.084	0.254	-0.025
	(0.042)	(0.028)	(0.046)	(0.032)	(0.068)	(0.052)
45 - 54	0.095	-0.153	0.146	-0.124	0.243	-0.063
	(0.040)	(0.024)	(0.044)	(0.027)	(0.067)	(0.046)
55 and above	0.004	-0.163	0.055	-0.141	0.183	-0.112
	(0.035)	(0.016)	(0.040)	(0.020)	(0.071)	(0.035)
Education						
has college	0.005	0.080	0.025	0.097	0.003	0.043
	(0.005)	(0.008)	(0.005)	(0.008)	(0.009)	(0.011)
Household annual income (yuan)						
$< 50,000$	-0.191	-0.150	-0.206	-0.166	-0.187	-0.167
	(0.017)	(0.022)	(0.016)	(0.020)	(0.020)	(0.024)
$[50,000; 100,000)$	-0.175	-0.106	-0.186	-0.123	-0.174	-0.128
	(0.024)	(0.031)	(0.023)	(0.029)	(0.031)	(0.037)
$[100,000; 150,000)$	-0.135	-0.076	-0.135	-0.083	-0.125	-0.080
	(0.020)	(0.031)	(0.020)	(0.030)	(0.030)	(0.038)
$[150,000; 200,000)$	-0.105	-0.035	-0.102	-0.041	-0.104	-0.037
	(0.022)	(0.036)	(0.022)	(0.035)	(0.031)	(0.044)
$[200,000; 250,000)$	-0.112	-0.029	-0.106	-0.032	-0.098	-0.029
	(0.021)	(0.039)	(0.022)	(0.038)	(0.032)	(0.047)
$[250,000; 300,000)$	-0.088	-0.007	-0.081	-0.007	-0.085	-0.020
	(0.026)	(0.043)	(0.027)	(0.043)	(0.036)	(0.050)
$\geq 300,000$	-0.086	0.048	-0.080	0.039	-0.072	0.046
	(0.028)	(0.056)	(0.030)	(0.055)	(0.043)	(0.069)
Commuting distance						
< 1 km	-0.164	-0.180	-0.143	-0.160	-0.187	-0.188
	(0.005)	(0.006)	(0.005)	(0.007)	(0.011)	(0.014)
$1 - 2.5$ km	-0.038	-0.050	0.013	-0.009	-0.057	-0.043
	(0.008)	(0.010)	(0.009)	(0.013)	(0.018)	(0.022)
$2.5 - 5$ km	0.044	0.051	0.116	0.120	0.032	0.080
	(0.009)	(0.013)	(0.011)	(0.016)	(0.023)	(0.029)
$5 - 10$ km	0.123	0.136	0.220	0.226	0.127	0.186
	(0.010)	(0.015)	(0.012)	(0.018)	(0.027)	(0.033)
$10 - 15$ km	0.147	0.179	0.254	0.280	0.151	0.268
	(0.012)	(0.017)	(0.014)	(0.020)	(0.029)	(0.036)
> 15 km	0.196	0.228	0.291	0.319	0.183	0.290
	(0.012)	(0.017)	(0.013)	(0.019)	(0.029)	(0.036)

cont'd on the next page

cont'd	model 1		model 2		model 3	
	male	female	male	female	male	female
Access to subway						
Home dist to nearest subway stn						
less than 0.5 km			-0.032	-0.007	-0.038	0.014
			(0.006)	(0.009)	(0.010)	(0.013)
0.5 - 1 km			-0.002	0.025	-0.007	0.029
			(0.007)	(0.011)	(0.012)	(0.015)
1 - 2 km			0.011	0.046	0.012	0.044
			(0.009)	(0.013)	(0.013)	(0.017)
more than 2 km			-	-	-	-
			-	-	-	-
Work dist to nearest subway stn						
less than 0.5 km			-0.126	-0.128	-0.135	-0.164
			(0.005)	(0.007)	(0.008)	(0.010)
0.5 - 1 km			-0.055	-0.062	-0.058	-0.083
			(0.007)	(0.009)	(0.010)	(0.011)
1 - 2 km			0.003	-0.033	-0.019	-0.067
			(0.009)	(0.011)	(0.013)	(0.013)
more than 2 km			-	-	-	-
			-	-	-	-
Spousal characteristics						
spouse has college degree					0.030	0.044
					(0.009)	(0.011)
spouse's log commute dist					-0.006	-0.018
					(0.003)	(0.004)
predicted % use car (mortgage)	50	33	61	51	60	57

Note: Sample from the 2015 Household Travel Survey. The dependent variable is an indicator whether the person drives or rides a car to work. All three models are estimated using a logit model. Each coefficient is associated with a variable that is the interaction between the male or female indicator and the indicated group. For example, the first coefficient reported in the table is associated with the male indicator interacted with an binary variable indicating the less than 34 year old group. The constant is suppressed. Marginal effects at the sample mean are reported with robust standard errors in parentheses.

B Additional Empirical Results

Table B.1: Gender Commute Gap in the United States

	(1)	(2)	(3)	(4)
	dep var: log commute time			
= 1 if male	0.103	0.112	0.106	0.108
	(0.001)	(0.001)	(0.001)	(0.001)
demographic controls		X	X	X
geographic controls			X	X
modes of transportation				X

Note: The data is from the 2017 American Community Survey, which represents 1% of the U.S. population. The sample includes individuals between 25 and 64 years old, not in group quarters, work at least half time (more than 20 weeks last year and more than 20 hours in a typical week), and commute to work (excluding those who work from home). Demographic controls include four age bins, whether having children under 6, marital status, and a set of detailed education attainment indicators. Geographic controls include the metropolitan statistical area (MSA) of residence. A state's rural area is also included as one "MSA." Variables on the mode of transportation include a list of indicators for major transportation modes.

Table B.2: Home Location Choice and the Gender Gap in Commutes, First-time Home Owners

	(1)	(2)	(3)	(4)
	dep var: ln dist btw work and new home			
= 1 if male	0.105	0.130	0.114	0.102
	(0.004)	(0.004)	(0.003)	(0.003)
Fixed effects				
age bin		X	X	X
detailed edu attain.		X	X	X
new home district			X	
household				X
N of obs	257402	257402	257402	257402

Note: Sample includes the households that do not already own a home. Robust standard errors are in parentheses.

Table B.3: Suppressed Coefficients from Table 4, Column 4

	husband's age group		
	under 30	31-40	41 and above
ln price for 1 sqm ×			
adj hhd inc in 1st quartile	-1.370	-1.007	-0.906
	(0.040)	(0.033)	(0.053)
adj hhd inc in 2nd quartile	-1.345	-1.082	-0.665
	(0.035)	(0.028)	(0.043)
adj hhd inc in 3rd quartile	-1.096	-0.820	-0.493
	(0.038)	(0.026)	(0.041)
adj hhd inc in 4th quartile	-0.576	-0.294	-0.231
	(0.040)	(0.023)	(0.035)
ln distance to current home ×			
adj hhd inc in 1st quartile	-1.159	-1.141	-1.111
	(0.015)	(0.012)	(0.020)
adj hhd inc in 2nd quartile	-1.072	-1.098	-1.038
	(0.013)	(0.010)	(0.016)
adj hhd inc in 3rd quartile	-1.038	-1.076	-0.950
	(0.014)	(0.009)	(0.015)
adj hhd inc in 4th quartile	-1.018	-1.036	-0.930
	(0.014)	(0.008)	(0.013)
= 1 if new community (omitted category)			
= 1 if community less than 5 years old ×			
adj hhd inc in 1st quartile	-1.428	-1.462	-1.223
	(0.047)	(0.038)	(0.054)
adj hhd inc in 2nd quartile	-1.508	-1.532	-1.427
	(0.042)	(0.032)	(0.044)
adj hhd inc in 3rd quartile	-1.574	-1.625	-1.508
	(0.046)	(0.031)	(0.043)
adj hhd inc in 4th quartile	-1.478	-1.553	-1.592
	(0.047)	(0.026)	(0.035)
= 1 if community between 6 and 10 years old ×			
adj hhd inc in 1st quartile	-2.329	-2.321	-2.298
	(0.051)	(0.041)	(0.065)
adj hhd inc in 2nd quartile	-2.505	-2.520	-2.569
	(0.044)	(0.034)	(0.052)
adj hhd inc in 3rd quartile	-2.649	-2.609	-2.857
	(0.048)	(0.032)	(0.050)
adj hhd inc in 4th quartile	-2.621	-2.617	-3.062
	(0.050)	(0.027)	(0.042)
= 1 if community between 11 and 15 years old ×			
adj hhd inc in 1st quartile	-2.269	-2.382	-2.406
	(0.054)	(0.044)	(0.075)
adj hhd inc in 2nd quartile	-2.509	-2.702	-3.004
	(0.046)	(0.037)	(0.062)
adj hhd inc in 3rd quartile	-2.739	-2.811	-3.189
	(0.050)	(0.034)	(0.058)
adj hhd inc in 4th quartile	-2.716	-2.829	-3.489
	(0.052)	(0.029)	(0.050)

Additional Coefficients from Table 4, Column 4 (*continued*)			
	husband's age group		
	under 30	31-40	41 and above
= 1 if community more than 16 years old ×			
adj hhd inc in 1st quartile	-2.558	-2.422	-2.711
	(0.043)	(0.040)	(0.069)
adj hhd inc in 2nd quartile	-2.771	-2.706	-3.193
	(0.046)	(0.034)	(0.057)
adj hhd inc in 3rd quartile	-3.032	-2.827	-3.468
	(0.049)	(0.031)	(0.054)
adj hhd inc in 4th quartile	-3.070	-3.090	-3.962
	(0.048)	(0.027)	(0.047)
community green area ratio ×			
adj hhd inc in 1st quartile	0.005	0.001	0.008
	(0.002)	(0.002)	(0.003)
adj hhd inc in 2nd quartile	0.006	0.006	0.012
	(0.002)	(0.001)	(0.002)
adj hhd inc in 3rd quartile	0.007	0.011	0.019
	(0.002)	(0.001)	(0.002)
adj hhd inc in 4th quartile	0.013	0.012	0.020
	(0.002)	(0.001)	(0.002)

Note: The table reports the suppressed coefficients and the corresponding standard errors of the interacted terms between community characteristics and 12 household type indicators in Column 4 of Table 4.

Table B.4: Home Location Choice and Gender Gap in Commuting, by Demographic Groups

	(1)	(2)	(3)	(4)	(5)	(6)
	dep var: ln commute dist btw work and new home					
= 1 if male	0.102	0.130	0.106	0.144	0.098	0.078
	(0.004)	(0.005)	(0.004)	(0.007)	(0.011)	(0.044)
Fixed effects						
age bin	X	X				
detailed edu attainment	X	X	X	X	X	X
home district	X	X	X	X	X	X
group	college	no college	< 35	$[35, 44]$	$[45, 54]$	$[55, 64]$
N of obs	146,244	151,919	184,211	79,036	31,659	3,257

Note: Robust standard errors in parentheses.

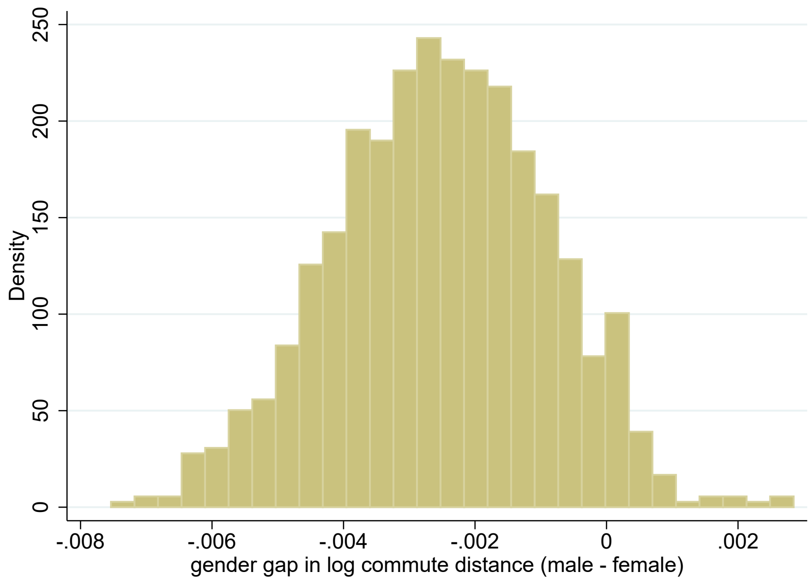
Table B.5: Commute Distance and Home Choice - Heterogeneity by Household Types

	(1) wife's age				(2) education $\{hus\ edu, wife\ edu\}$		
	ln comm dist				ln comm dist		
	husband	wife	β^f/β^m		husband	wife	β^f/β^m
$\notin [30,39]$	-0.207	-0.343	1.661	$\{H,H\}$	-0.226	-0.416	1.839
	(0.007)	(0.007)	[0.000]		(0.008)	(0.008)	[0.000]
$\in [30,39]$	-0.260	-0.439	1.690	$\{H,L\}$	-0.246	-0.282	1.147
	(0.007)	(0.007)	[0.000]		(0.014)	(0.013)	[0.144]
				$\{L,H\}$	-0.181	-0.428	2.362
					(0.016)	(0.015)	[0.000]
				$\{L,L\}$	-0.245	-0.389	1.590
					(0.008)	(0.008)	[0.000]

	(3) income difference, quartile $\ln(hus\ inc) - \ln\{wife\ inc\}$				(4) age difference $hus\ age - wife\ age$		
	ln comm dist				ln comm dist		
	husband	wife	β^f/β^m		husband	wife	β^f/β^m
1^{st}	-0.204	-0.421	2.062	< 0	-0.237	-0.367	1.549
	(0.009)	(0.009)	[0.000]		(0.012)	(0.011)	[0.000]
2^{nd}	-0.244	-0.444	1.822	$[0,1]$	-0.243	-0.391	1.607
	(0.012)	(0.011)	[0.000]		(0.008)	(0.008)	[0.000]
3^{rd}	-0.201	-0.363	1.808	$[2,3]$	-0.208	-0.414	1.990
	(0.011)	(0.011)	[0.000]		(0.010)	(0.010)	[0.000]
4^{th}	-0.262	-0.342	1.306	≥ 4	-0.226	-0.378	1.668
	(0.009)	(0.009)	[0.000]		(0.012)	(0.012)	[0.000]

Note: Each column shows the coefficients associated with the husband's and the wife's log commute distances interacted with variables capturing household division of labor and within-household bargaining power. The ratio between the coefficient associated with the wife's commute distance (β^f) and that with the husband's commute distance (β^m) is reported with $p-$value from testing $\beta^f/\beta^m = 1$ in brackets. The rest of the specification is the same as Column 4 of Table 4. Coefficients associated with 12 household types are suppressed.

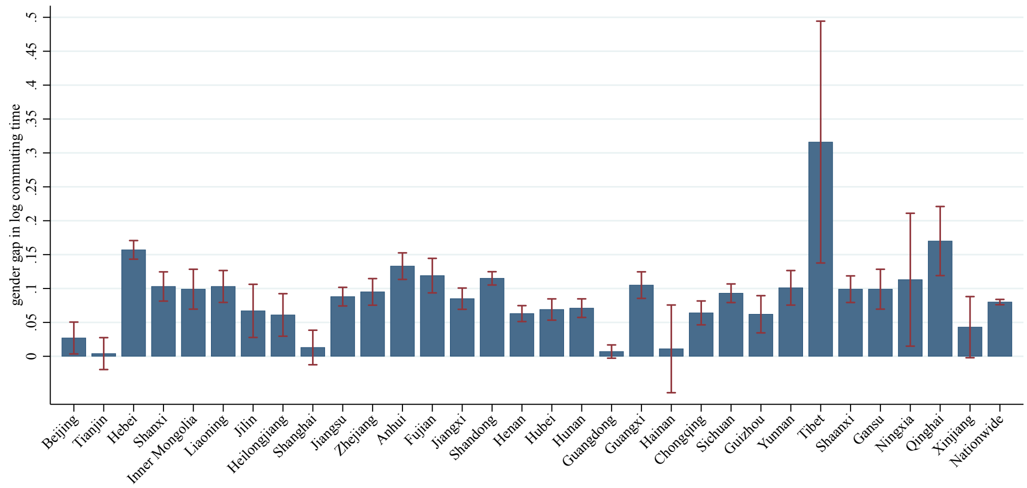
Figure B.1: Distribution of Gender Commute Gap Using Placebo New Homes



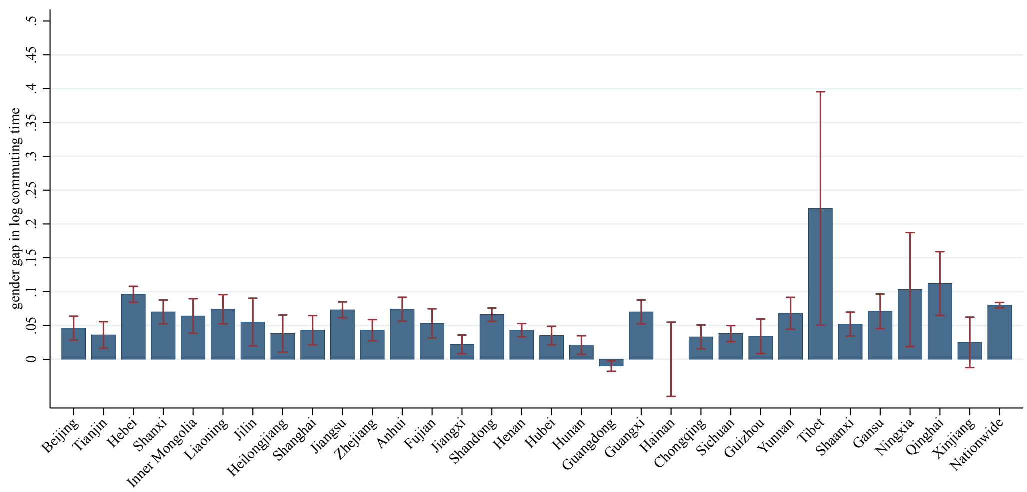
Note: Density graph of the coefficient associated with a male indicator in placebo tests using the same specification as in Column 4 of Table 3. For each placebo test, new homes in the sample are reshuffled and randomly assigned to a household. We repeat 1000 placebo estimates. Estimation of the original sample has a coefficient of 0.105, which is well outside of the distribution of placebo effects.

Figure B.2: Gender Commute Gap by Province

Panel A



Panel B



Note: The data is from the 2015 inter-decennial population census. The sample includes individuals between 18 and 64 year old who are currently employed in non-agricultural industries. Each bar presents the estimated coefficient associated with an indicator for men in each province. In Panel A, regressions include a set of age bin fixed effects and a set of education attainment fixed effect. In Panel B, regressions include an additional set of transportation mode fixed effects. Range bars indicate the t95% confidence intervals.