Productivity Externalities of Working from Home: Welfare and Policy Implications *

Huilin Zhang

June 23, 2025

Please click here for the latest version

Abstract

I study the socially optimal mix of onsite work and remote work. I develop a general equilibrium model in which workers decide how much to work onsite and work from home. Productivity spillovers can occur within and between onsite and remote workers. The model predicts that the balance between onsite and remote productivity spillover effects affects the gap between the socially optimal and the market equilibrium level of onsite work. I measure these spillovers by matching the model to U.S. survey data from 2022 to 2024 at the city-sector-work mode level. I find that, on average, a social planner could improve welfare by 2% by increasing hybrid workers' share of onsite time by 3% and increasing the number of fully onsite workers by 2%. This could be accomplished by offering a subsidy for onsite work equal to 11% of hybrid workers' gross income. Without the remote productivity spillovers, a similar level of welfare improvement would require larger changes: hybrid workers' share of onsite time would need to increase by 5%, and the number of fully onsite workers would need to increase by 3%. The subsidy would cost 15% of hybrid workers' gross income.

Keywords: productivity externality, remote collaboration technology, working from home, social welfare, optimal policy

JEL Classication: C54, D62, H23, R13

* I am extremely grateful to Farid Farrokhi, Chong Xiang, David Hummels, and Victoria Prowse for their invaluable support and guidance on this project. I am also thankful to Yichen Su, Lindsay Relihan, Sunham Kim, John R Umbeck, J Ryan Umbeck for their insightful feedback.

1 Introduction

We have witnessed a substantial increase in working from home (WFH) due to the COVID-19 pandemic. In 2024, the average worker in the U.S. spent around 28% of workdays WFH, which is a 4-fold increase compared to the pre-pandemic level (around 7%) (Barrero et al. (2021)). The widespread adoption of WFH can be attributed to workers' appreciation of its flexibility (Barrero et al. (2021)) and improvements in remote collaboration technology such as video conferencing. However, there are concerns that less in-person interaction reduces productivity since the productivity spillover from physical proximity has been well-documented in the literature. Is the shift to WFH socially optimal? One key to answering this question lies in whether new technologies generate positive externalities among remote workers that might compensate for reduced externalities from in-person work. The stronger the remote productivity spillover, the more likely it is that workers enjoy the amenity of working from home without reducing overall productivity.

While the literature has separately studied productivity spillovers from in-person work and the benefits of remote collaboration, these papers have not addressed the relative strength of onsite and remote productivity spillovers and how these spillovers affect the social optimum. This paper quantifies onsite and remote productivity spillovers, the gap between the socially optimal and the market equilibrium level of onsite work, and the optimal policies to achieve the social optimum.

I begin by documenting two empirical facts about onsite and remote work: first, the extent of remote work varies by city and sector, and second, remote (onsite) workers' residual wages are positively correlated with remote (onsite) employment size across cities and sectors.

Motivated by the empirical pattern, I employ a Roy-Fréchet model in which workers in each city choose a sector and one of three work modes (onsite, hybrid, fully WFH). Hybrid workers choose their shares of time spent working onsite and WFH by comparing the productivity and disutility of working onsite versus WFH. Productivity spillovers can occur within and between groups of onsite labor and remote labor. Similar to previous work (Davis et al. (2024)), an onsite (remote) worker will be more productive when there are more onsite (remote) workers and hybrid workers spend more time working onsite (WFH). In contrast to previous work, the model allows cross-worksite spillovers: more onsite labor can increase a remote worker's productivity, and vice versa.

The cross-worksite spillover parameter governs how much cross-worksite labor contributes to productivity relative to within-worksite labor. Work arrangements between onsite and remote do not affect aggregate productivity when cross-worksite and within-worksite productivity spillovers are symmetric. Otherwise, I measure the effect of work arrangements on productivity at the extensive and intensive margins. The extensive margin is measured by the elasticity of onsite (remote) productivity with respect to the total labor contributing to it, which I refer to as the onsite (remote) agglomeration elasticity. The intensive margin is measured by the elasticity of onsite (remote) productivity with respect to hybrid workers' share of onsite (remote) time in each city-sector cell, which I refer to as the onsite (remote) time elasticity. A larger gap between onsite and remote time elasticities leads to a larger gap between the social optimum and the market equilibrium for hybrid workers' onsite share. I'm not aware of any literature that directly investigates this intensive margin of productivity spillover.

I match the model to the U.S. 2022–2024 Current Population Survey (CPS) and the 2022 American Community Surveys (ACS) data for estimation and quantification. From the CPS data, I obtain average employment, residual wages, and the share of time spent working onsite at the core-based statistical area (CBSA), industry, and work mode levels. From the ACS data, I calculate the average commuting time by city, which is used to measure onsite disutility. I use these observable variables to recover the model shifters and structurally estimate the agglomeration elasticities and the cross-worksite spillover.

Inspired by Ahlfeldt et al. (2015) and Farrokhi (2021), the estimation applies the generalized method of moments (GMM) and uses model shifters to construct moment conditions. The estimation process consists of an inner loop and an outer loop. In the inner loop, given the values of parameters and observed data, I obtain labor supply and demand shifters and WFH disutility using the model inversion method (Redding and Rossi-Hansberg (2017)). In the outer loop, I obtain the estimates that minimize the sum of squared errors of the moments. The moment conditions are similar to identification assumptions in instrumental variable (IV) regressions. They jointly identify parameters, but each is more closely related to a specific parameter. The identification assumption for remote agglomeration elasticity is that the remote labor supply shifter (which resembles an IV for remote labor) is uncorrelated with the exogenous remote labor demand shifter. The identification assumption for cross-worksite spillover is that the relative onsite disutility (which resembles an IV for the share of time spent working onsite) is uncorrelated with the exogenous relative onsite productivity.

The estimated onsite agglomeration elasticity is larger than the remote one, indicating a stronger onsite productivity externality at the extensive margin. The cross-worksite spillover is small, implying a notable efficiency loss for the productivity spillover between onsite and remote work. Based on these three estimates and the share of hybrid workers' onsite (remote) time in total labor across city-sector cells, I calculate the onsite (remote) time elasticities.

I then measure the gap between the social optimum and the market outcome and calculate the subsidies required to achieve the social optimum. The subsidies incentivize hybrid workers to adjust their shares of time spent working onsite to achieve the socially optimal levels of onsite time and employment composition through spillover and labor mobility effects.

The main findings are that, on average, the social optimum favors more onsite work, but the existence of remote productivity spillovers means that the market equilibrium is closer to the socially optimal level of onsite work than in a case with no remote spillovers. On average, a social planner could improve welfare by 2% by increasing hybrid workers' share of onsite time by 3% and increasing the number of fully onsite workers by 2%. This could be accomplished by levying an income tax and offering a subsidy to onsite work equal to 11% of hybrid workers' gross income. Without remote productivity spillovers, a similar level of welfare improvement would require larger changes: hybrid workers' share of onsite time would need to increase by 5%, and the number of fully onsite workers would need to increase by 3%. Achieving this outcome would require a higher subsidy—15% of hybrid workers' gross income when funded by an income tax.

Hybrid workers' time allocation between onsite and remote work, and the welfare losses resulting from it, depends on the strength of productivity spillovers and the amenity values of each option. Because these factors vary across cities and sectors, so does the departure of the market equilibrium from the social optimum. The social optimum tends to have less onsite time for hybrid workers than the market equilibrium in city-sector cells where (1) the intensive onsite productivity spillover (measured by the onsite time elasticity) is weaker than the remote one, and (2) commuting costs are higher than the WFH disutility. The gap between the socially optimal and the market equilibrium level of hybrid workers' onsite share tends to widen as the gap between onsite and remote time elasticities increases. The socially optimal employment allocation features a shift from initially high-income to low-income sectors and results in a reduction of the sectoral income premium for the three work modes.

This paper relates to four strands of literature. The first strand of literature empirically studies how remote work affects productivity spillovers. Some researchers document the negative effect of WFH on onsite productivity spillovers: Frakes and Wasserman (2021) and Emanuel et al. (2022) show that increased WFH reduces knowledge sharing at the individual level using citation patterns and online feedback data, respectively. Liu and Su (2022) show that more WFH reduces the urban agglomeration effect and urban wage premium. Other researchers document the benefits of remote collaboration at the firm (Forman and Van Zeebroeck (2012) and Forman and van Zeebroeck (2019)), institution (Presidente and Frey (2023)), and team levels (Brucks and Levav (2022)).

The second strand of literature uses structural models to study the effect of remote work. Researchers make different assumptions about productivity spillover effects. Some studies assume that remote work does not contribute to productivity spillovers (Safirova (2002); Rhee (2008); Behrens et al. (2021); Delventhal et al. (2022); Monte et al. (2023)), while others consider the possibility that remote work generates productivity spillovers

(Lennox (2020), Davis et al. (2024), Delventhal and Parkhomenko (2023)). The model in this paper features endogenous time allocation and imperfect substitution between working onsite and WFH as in Delventhal and Parkhomenko (2023) and Davis et al. (2024). This paper complements the existing literature by modeling and measuring productivity spillovers within and across onsite and remote work. Additionally, I use post-pandemic data to study the socially optimal level of onsite work, whereas previous studies match pre-pandemic data with their models to predict the effects of the WFH shock and use post-pandemic data to verify those model predictions.

Third, this paper relates to the literature studying WFH policies. Safirova (2002) examines the subsidies and taxes required to achieve the socially optimal land allocation in a city with remote workers. Bertram et al. (2024) use a principal-agent model to study the optimal level of remote work and the related incentive scheme in a firm. Ding and Ma (2023) and Flynn et al. (2024) document firms' return-to-office policies. This paper considers socially optimal policies and quantifies city-sector-specific subsidies that incentivize hybrid workers to adjust their onsite work time.

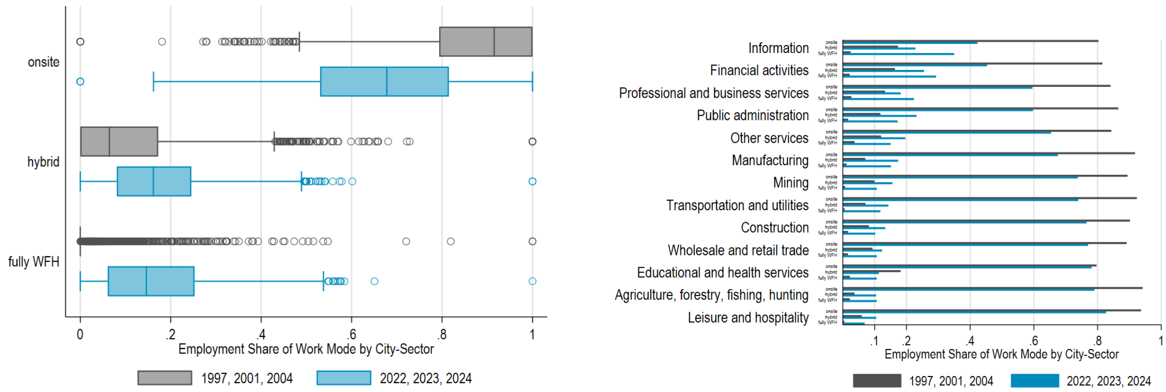
Finally, this paper is related to the literature that studies the socially optimal allocation. in the presence of productivity spillovers across workers (Fajgelbaum and Gaubert (2020); Rossi-Hansberg et al. (2019)). In these studies, market equilibrium deviates from social optimum in terms of employment. This paper shows that the market equilibrium may also deviate from the social optimum in terms of work time allocation between working onsite and WFH.

The structure of the paper is as follows. Section 2 discusses empirical facts about onsite and remote work. Section 3 presents a basic theoretical model with homogeneous workers. Section 4 presents the full model with variations in cities, sectors, and work modes. Section 5 describes the data and explains the structural estimation of the parameters. Section 6 analyzes the quantification results for the gap between the market outcome and the social optimum and the socially optimal subsidies. Section 7 concludes.

2 Empirical Patterns of Working From Home

This section documents the increase in WFH, the variations of remote work across cities and sectors, the evidence suggesting remote externalities, and the residential and workplace patterns of onsite and remote workers.

The COVID-19 pandemic has caused a significant shift to WFH in the US. according to Barrero et al. (2023), the average share of full paid days WFH gradually increased from around 3% in 2003 to 7.2 % in 2019. After the pandemic, the average days WFH have persisted at around 30 % from 2022 to 2024 (see Figure 16 in the appendix).



(a) Distribution of Work Mode Composition Across City-Sector Cells

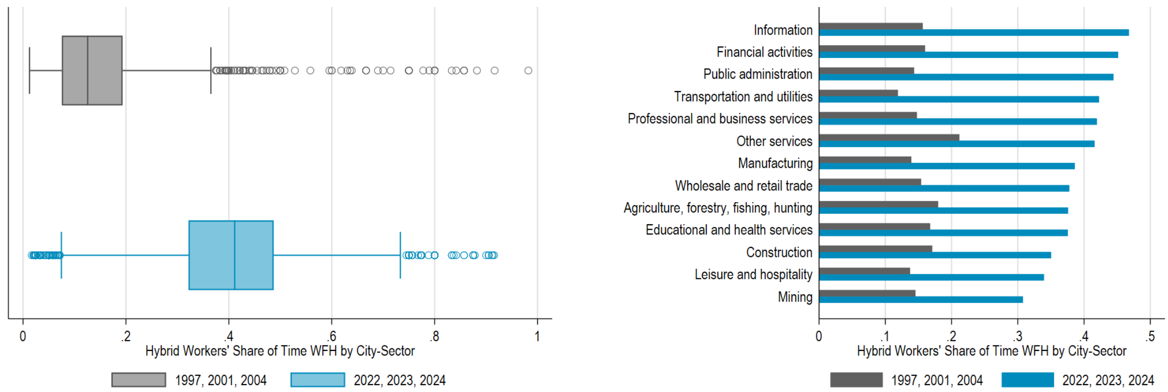
(b) City Averages of Employment Share by Work Mode and Sector

Figure 1: Increased Hybrid and Fully Working from Home Workers

Note: Data source is Current Population Survey (CPS) work schedule supplement at 1997, 2001, and 2004 May and basic monthly survey from 2022 October to 2024 December. The samples are labor forces excluding self-employed, armed forces, and unpaid family workers. The employment share of onsite, hybrid, or fully WFH workers by city-sector is the percentage of these workers relative to total employment in a city-sector cell.

The increase in WFH also appears at the city-sector level in the margins of employment share (Figure 1) and work time (Figure 2). Figure 1(a) compares the distribution of the share of onsite, hybrid, and fully WFH workers at the city-sector level before and after the WFH shock. It shows that the shares of hybrid and fully WFH workers increase in the first three quantiles of the distribution. The median of onsite shares drops from around 90 % in 1997, 2001, and 2004 to around 70 % in 2022 to 2024. Figure 1(b) compares the city averages of onsite, hybrid, and fully WFH shares in 13 sectors before and after the WFH shock. Across all sectors, the average shares of hybrid and fully

WFH shares increase, and the onsite shares decrease. In summary, Figure 1 displays an increase in hybrid and fully WFH workers across cities and sectors. Figure 2(a) shows the increase in hybrid workers' share of time spent WFH across city-sector for the first three quantiles. Figure 2(b) shows that across all sectors, the city average of the share of time spent WFH for hybrid workers increases.



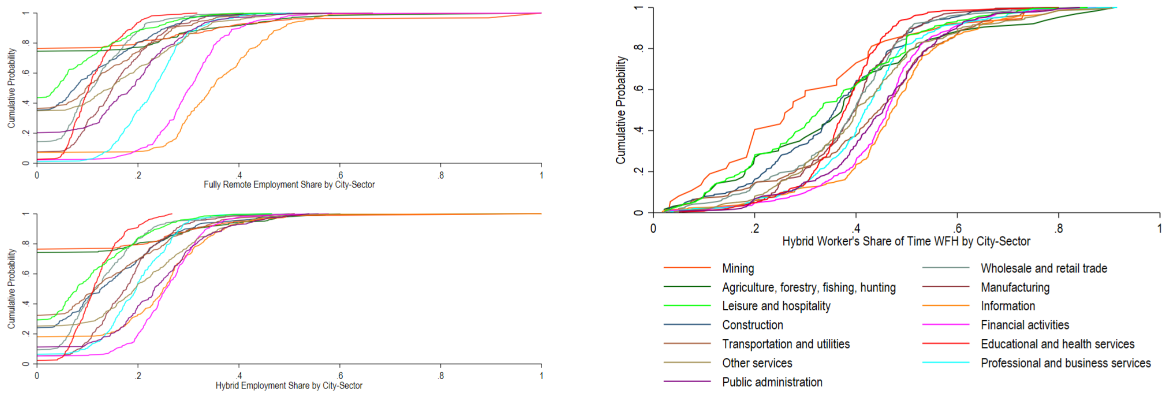
(a) Distribution of the Average Share of Time WFH by Hybrid Workers Across City-Sector Cells

(b) City Averages of Hybrid Workers' Share of Time WFH by Sector

Figure 2: Hybrid Workers Spend More Time Working from Home

Note: Data source is CPS work schedule supplement at 1997, 2001, and 2004 May and basic monthly survey from 2022 October to 2024 December. The samples are labor forces excluding self-employed, armed forces, and unpaid family workers. The share of time spent WFH is calculated as the number of hours WFH divided by the total number of hours worked in a week.

Focusing on the measures after the WFH shock, the level of remote work varies across cities and sectors. Figures 1(b) and 2(b) demonstrate the sectoral variations in remote work after the WFH shock. Based on the average across cities, information and financial activities have over 20% hybrid or fully WFH workers, while agriculture and hospitality have less than 10%. Hybrid workers in the information and financial activities sectors spend nearly 50% of their work time WFH, whereas those in mining spend around 30% of their work time WFH. Figure 3 shows the cumulative distribution of remote employment or work time shares by sector across cities. Different curves show the variations in remote work levels across sectors. Each curve in Figure 3(a) shows that within a sector, fully WFH and hybrid employment shares vary across cities. Figure 3(b) shows that within each sector, hybrid workers' WFH time varies across cities.



(a) Cumulative Distribution of Employment Share (b) Cumulative Distribution of Hybrid Workers' Share of Time WFH

Figure 3: Variations of Remote Employment and WFH Time Across Cities and Sectors
Note: Data source is CPS basic monthly survey from 2022 October to 2024 December. The samples are labor forces excluding self-employed, armed forces, and unpaid family workers.

With the rise of WFH, the methods of knowledge sharing have shifted from relying primarily on in-person meetings to an increasing use of virtual interactions thanks to the widely adoption of video conferencing. Telecommunication technology facilitates knowledge sharing among remote workers and between onsite and remote workers. Figure (4) shows that a higher number of onsite, hybrid, or fully WFH workers is associated with higher average residual wages for these worker groups across city-sector cells. The positive correlation suggests that productivity externalities exist not only among onsite workers but also among remote workers.

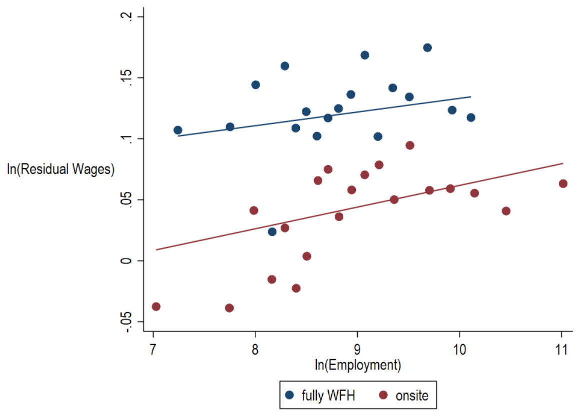


Figure 4: Onsite and Remote Productivity Externalities at the City-Sector Level
Note: Data source is CPS basic monthly survey from 2022 October to 2024 December. The samples are labor forces excluding self-employed, armed forces, and unpaid family workers. The figures are binned scatterplots of residual wages against employment with the city fixed effect. Appendix H describes the method for obtaining residual wages.

Figure 4 implies the presence of onsite and remote productivity spillover within cities. The onsite agglomeration effect within a city has been well documented. One may wonder to what extent the remote productivity externality within cities align with real-world practice, given that WFH provides feasibility in terms of work location. Table 1 reveals that nearly 90% of onsite workers and 98% of home-based workers live and work in the same CBSA, both before and after the WFH shock. Additionally, Table 5 in the appendix shows that over 80% of onsite and home-based workers do not move across CBSAs in the short term. These empirical facts suggest local workforce stability, which may stem from two factors: first, the prevalence of onsite and hybrid work arrangements that encourage workers to live within commuting distance, and second, the influence of agglomeration effects on fully WFH workers. For example, remote workers live in the workplace cities to maintain professional connections within their local sector. This local workforce stability for remote workers contributes to the observed remote productivity externality at the CBSA-sector level.

Motivated by these empirical facts, I build a model incorporating onsite and remote productivity externalities. I begin with a basic model that focuses on implications of the productivity spillover pattern within one city with homogenous workers. In the full model, I introduce the variations by city, sector, and work mode.

Table 1: Most Onsite and Home-based Workers Live and Work in the Same CBSA

	2018-2019		2021-2023	
	on-site worker	home-based worker	on-site worker	home-based worker
Work and live in the same CBSA (%)	88.55	97.83	87.96	98.21
Work and live in different CBSA (%)	11.45	2.17	12.04	1.79
N (unweighted)	2,644,257		4,016,093	

Note: Data source is American Community Survey (ACS). Samples are wage workers excluding unpaid family workers and military workers. Each observation's residence or workplace (PUMA or PWPUMA) is matched to one CBSA with the largest population share. 5.94 % of the samples cannot be classified as either home-based or onsite workers due to nonresponse regarding their commuting mode.

3 The Basic Model

3.1 The Market Outcome

3.1.1 Preference and Production

Consider a city with $\bar{L}$ homogenous workers. Each worker chooses the share of time spent working onsite and consumes tradable goods and housing. The utility function of a worker is:

$$U = \frac{c^{\alpha} \hbar^{1-\alpha}}{\bar{\phi} d(\theta)}, \tag{1}$$

where c and h are the quantity of tradable goods and housing, respectively. The constant $\bar{\phi} = \alpha^{\alpha}(1-\alpha)^{1-\alpha}$. $\theta \in [0, 1]$ is the share of time spent working onsite. With total work time normalized to 1, the remaining portion, $1 - \theta$, represents the share of time spent WFH.

The disutility $d(\theta)$ is a weighted average of disutility for onsite work (ζ^o)[1] and amenity costs of WFH (ζ)[2]:

$$d(\theta) = \theta\zeta^o + (1 - \theta)\zeta. \tag{2}$$

The derivative of disutility with respect to the share of time spent working onsite is:

$$d'(\theta) = \zeta^o - \zeta. \tag{3}$$

When onsite disutility equals WFH amenity costs, total disutility is a constant. When onsite disutility is greater than WFH amenity cost ($\zeta^o > \zeta$), total disutility increases with the share of time spent working onsite; conversely, when onsite disutility is less than WFH amenity cost, total disutility decreases with the share of time spent working onsite.

The budget constraint of the worker is $Pc+qh = I(\theta)$, where P is the price of tradable goods, q denotes the housing price, and $I(\theta)$ refers to income. Income combines wages ($W(\theta)$) and returns from returns to land (R). Assuming the ratio of land returns to wages

[1] In the full model, I use commuting costs to measure onsite disutility.

[2] This function form follows Delventhal and Parkhomenko (2023). It leads to a concise and intuitive solution of θ in market equilibrium (7).

is $r = R/W(\theta)$, then workers' income can be expressed as $I(\theta) = (1+r)W(\theta)$. A worker's total wage is the unit wage multiplied by the efficiency unit of labor: $W(\theta) = w\ell(\theta)$. The efficiency unit of labor $\ell(\theta)$ is determined by the share of time spent working onsite, remote productivity, and onsite productivity:

$$\ell(\theta) = [(A(1-\theta))^{\frac{\rho-1}{\rho}} + (B\theta)^{\frac{\rho-1}{\rho}}]^{\frac{\rho}{\rho-1}}, \tag{4}$$

where remote productivity (A) and onsite productivity (B) are given constants from workers' perspective. ρ denotes the elasticity of substitution of WFH and working onsite.

Given the share of time spent working onsite, the indirect utility function is:

$$u = \frac{(1+r)W(\theta)}{d(\theta)P^{\alpha}q^{1-\alpha}}. \tag{5}$$

Workers face trade-offs between onsite-WFH amenities and onsite-remote productivity when deciding the share of time spent working onsite. Spending more time working onsite means (1) experiencing more disutility from onsite work but less amenity cost from WFH, and (2) relying more on onsite productivity and less on remote productivity to produce outputs and earn wages. Maximizing the indirect utility with respect to θ yields the first-order condition that characterizes the share of time spent working onsite chosen by workers [3]:

$$\frac{\theta}{1-\theta} = \left(\frac{B}{A}\right)^{\rho-1}\left(\frac{\zeta}{\zeta^{o}}\right)^{\rho}. \tag{6}$$

Equation (6) indicates that the relative onsite share increases with relative onsite productivity and relative onsite amenity (inverse of onsite disutility). Rearranging this equation gives the solution for the share of time spent working onsite:

$$\theta = \frac{1}{1 + (\frac{\zeta^{o}}{\zeta})^{\rho}(\frac{A}{B})^{\rho-1}}. \tag{7}$$

The production function of tradable goods is $Y = \ell(\theta)L$, where L is the number of workers. By profit maximization, wage is $W(\theta) = P\ell(\theta)$. The price of tradable goods P

[3] Appendix B shows the derivation process.

is normalized to 1.

The construction firm uses land and tradable goods to produce housing. Assume the production function for the housing sector is

$$H = K^{\gamma}\bar{H}^{1-\gamma}, \tag{8}$$

where K is the tradable goods used for producing housing. $\bar{H}$ is land used. γ is the share of tradable goods used in producing housing. The housing supply elasticity is $\frac{\gamma}{1-\gamma}$, thus $\gamma = 0$ suggests a perfectly inelastic housing supply; $\gamma = 1$ corresponds to a perfectly elastic housing supply.

The profit maximization problem of the housing production firm is $\max_{K,\bar{H}} \pi = qK^{\gamma}\bar{H}^{1-\gamma} - pK - r\bar{H}$, where r denotes the unit returns to land. Solving the problem gives the total land returns: $RL = r\bar{H} = (1-\gamma)qH$. Combining housing market clearing ($hL = H$) and demand for housing ($h = (1-\alpha)\frac{I}{q}$) gives the housing quantity $H = \frac{(1-\alpha)IL}{q}$. The returns to land becomes $RL = (1-\alpha)(1-\gamma)IL = (1-\alpha)(1-\gamma)(W(\theta)+R)L$. Rearranging this equation gives the ratio of rent to wages as $\bar{r} = \frac{1}{\alpha+\gamma(1-\alpha)} - 1$. Therefore, income can be expressed as $I = \frac{1}{\alpha+\gamma(1-\alpha)}W(\theta)$.

By the profit maximization problem of the housing production firm, the demand for tradable goods used to produce housing is $K = \gamma\frac{qH}{P}$. Substituting the housing demand $H = \frac{1-\alpha}{q}I(\theta)L$ into this equation, K can be expressed as a function of income: $K = \gamma(1-\alpha)\frac{I(\theta)L}{P}$. Combining this equation with housing market clearing, demand for housing, and the housing production function yields the housing price as $q = \left(\frac{(1-\alpha)I(\theta)L}{\bar{H}}\right)^{1-\gamma}\left(\frac{P}{\gamma}\right)^{\gamma}$.

3.1.2 The Extensive Margins of Onsite and Remote Externalities

The aggregate remote productivity increase in the total remote labor and the contribution from onsite labor:

$$A = \bar{a}[(1-\theta)L + \tau\theta L]^{\lambda^R}. \tag{9}$$

The aggregate onsite productivity increase in the total onsite labor and the contribution from remote labor:

$$B = \bar{b}[\theta L + \tau(1-\theta)L]^{\lambda}. \tag{10}$$

$\bar{a}, \bar{b}$ are exogenous remote and onsite productivity, respectively. λ^R and λ are elasticities of remote and onsite productivity with respect to aggregate labor, respectively. They govern the strength of the extensive margin of remote and onsite productivity externalities. If $\lambda^R = 0$, remote productivity is exogenous.[4]

$\tau \in [0, 1]$ denotes cross-worksite spillover. $\tau = 0$ means no cross-worksite spillover. Remote work only contributes to remote productivity and onsite labor only contributes to onsite productivity.[5] $\tau > 0$ indicates the presence of productivity spillover between onsite and remote workers.

$\tau = 1$ means no efficiency loss in cross-worksite spillover. Onsite labor contributes to remote productivity in the same way as remote labor, and remote labor contributes to onsite productivity in the same way as onsite labor. In this case, aggregate productivity becomes $A = \bar{a}L^{\lambda^R}$ and $B = \bar{b}L^{\lambda}$, which is irrelevant to the allocation of time between onsite work or WFH.

$\tau \in (0, 1)$ implies an efficiency loss in cross-worksite spillover. The contribution of onsite labor to remote productivity is smaller than that of remote labor, and the contribution of remote labor to onsite productivity is smaller than that of onsite labor. This may occur due to information friction during telecommunication.[6]

If all workers work onsite, the efficient labor unit is only determined by onsite productivity ($\theta = 1$,$\ell(\theta = 1) = B = \bar{b}L^{\lambda}$).

[4] The scenario where $\lambda^R = 0$ and τ is either 0 or 1 is similar to the model in Delventhal and Parkhomenko (2023).

[5] Davis et al. (2024)'s model considers this separation of productivity spillover.

[6] Imagine an onsite speech that is broadcast live on the Internet. Both onsite and remote audiences can ask the speakers questions. When online audiences ask questions, the speaker relies on the Internet to receive the questions in text, video, or audio form, but the communication between onsite speakers and onsite audiences does not have such a constraint. As a result, remote labor contributes less to onsite productivity than onsite labor. For remote productivity, remote audiences may lose the sound from the speakers or onsite discussion if there are technical issues with the microphone. On the other hand, remote audiences may discuss the speech in the chat box, which is more manageable compared to a discussion between remote audiences and onsite people because remote audiences use the same form of communication. Because communication between remote audiences is smoother than receiving information from onsite people, onsite labor contributes less to remote productivity than remote labor.

3.1.3 The Share of Time Spent Working Onsite in Market Equilibrium

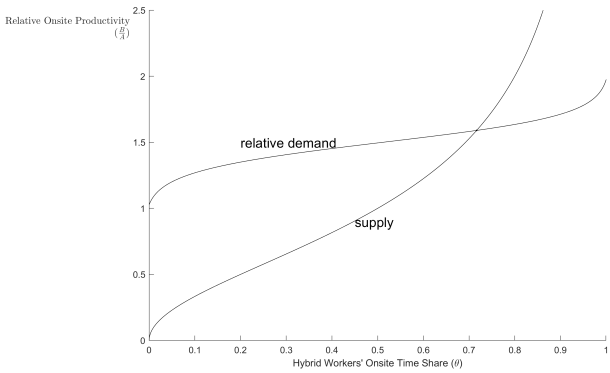


Figure 5: Share of Time Spent Working Onsite and Relative Onsite Productivity

This section discusses how the share of time spent working onsite (θ) is determined in the market equilibrium. Equation (7) shows that workers choose the share of time working onsite based on relative onsite productivity and it characterizes the supply side of the onsite time:

$$\theta = f(\frac{B}{A}) = \frac{1}{1 + (\frac{\zeta}{\zeta^o})^{-\rho}(\frac{B}{A})^{1-\rho}}.$$

Equations (9) and (10) imply that the relative onsite productivity is a function of the share of time spent working onsite:

$$\frac{B}{A} = f(\theta) = \frac{\bar{b}(\tau(1-\theta)L + \theta L)^{\lambda}}{\bar{a}[(1-\theta)L + \tau\theta L]^{\lambda^R}} = \frac{\bar{b}[\tau(1-\theta) + \theta]^{\lambda}}{\bar{a}[(1-\theta) + \tau\theta]^{\lambda^R}} L^{\lambda - \lambda^R}.$$

This equation characterizes the relative demand curve of the onsite time. If cross-worksite spillover $\tau \in [0, 1)$, the relative onsite productivity increases with the share of time spent working onsite and the relative demand curve is upward sloping. If cross-worksite spillover $\tau = 1$, the relative demand curve is irrelevant with the value of θ and becomes a horizontal line.

The share of time spent working onsite and relative onsite productivity are simultaneously determined. The intersection of the relative demand and supply curve for onsite time pins down the market outcome of θ as shown in Figure 5.

3.2 The Socially Optimal Share of Time Spent Working Onsite

3.2.1 The First Order Condition and the Intensive Margins of Productivity Externalities

In contrast to market equilibrium, a social planner will internalize the productivity externalities when choosing the share of time spent working onsite. Appendix C shows the process of solving the social planner problem. The first-order condition characterizing the socially optimal share of time spent working onsite is

$$\frac{\theta}{1-\theta} = \left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1} \left[\frac{\zeta + d(\theta)\left(\tilde{\gamma}\left(\overbrace{\delta}^{\text{onsite productivity externality}}+1\right)-1\right)}{\zeta^o + \underbrace{d(\theta)}_{\substack{\text{composite of}\\ \text{disutility}}}\left(\underbrace{\tilde{\gamma}}_{\substack{\text{housing}\\ \text{congestion}}}\left(\underbrace{\delta^R}_{\substack{\text{remote}\\ \text{productivity}\\ \text{externality}}}+1\right)-1\right)}\right]^{\rho}, \tag{11}$$

where $A(\theta) = \bar{a}((1-\theta)\bar{L} + \tau\theta\bar{L})^{\lambda^R}$, $B(\theta) = \bar{b}(\theta\bar{L} + \tau(1-\theta)\bar{L})^{\lambda}$. $\tilde{\gamma} = \alpha + \gamma(1-\alpha)$. $d(\theta) = \theta\zeta^o + (1-\theta)\zeta$.

δ and δ^R denote the intensive margin of onsite and remote productivity externalities, respectively:

$$\delta = \frac{\partial B(\theta)/B(\theta)}{\partial\theta/\theta} = \lambda(1-\tau)\frac{\theta\bar{L}}{\theta\bar{L} + \tau(1-\theta)\bar{L}}. \tag{12}$$

$$\delta^R = \frac{\partial A(\theta)/A(\theta)}{\partial(1-\theta)/(1-\theta)} = \lambda^R(1-\tau)\frac{(1-\theta)\bar{L}}{(1-\theta)\bar{L} + \tau\theta\bar{L}}. \tag{13}$$

δ refers to the elasticity of onsite productivity with respect to the share of time spent working onsite. δ^R refers to the elasticity of remote productivity with respect to the share of time spent WFH. The onsite time elasticity (δ) equals the onsite agglomeration elasticity (λ) multiplied by the cross-worksite efficiency loss $(1-\tau)$ and the share of onsite labor in aggregate labor that contributes to the onsite productivity ($\frac{\theta\bar{L}}{\theta\bar{L}+\tau(1-\theta)\bar{L}}$). Similarly, the remote time elasticity (δ^R) equals the remote agglomeration elasticity (λ^R) multiplied by the cross-worksite efficiency loss $(1-\tau)$ and the share of remote labor in aggregate labor that contributes to the remote productivity ($\frac{(1-\theta)\bar{L}}{(1-\theta)\bar{L}+\tau\theta\bar{L}}$).

Note that the onsite (remote) time elasticity may differ from extensive margin of onsite (remote) productivity externality ($\delta \leq \lambda, \delta^R \leq \lambda^R$). When productivity spillover exists between onsite and remote workers and efficiency losses for onsite-remote interaction ($\tau \in (0,1)$), the onsite (remote) time elasticity is smaller than the extensive margin of onsite (remote) productivity externality. In contrast, when there is no cross-worksite spillover ($\tau = 0$), the onsite (remote) time elasticity equals the extensive margin of onsite (remote) productivity externality in this basic model with homogenous workers. In the model where not all workers have the flexibility to choose working onsite or WFH, the absence of cross-worksite spillover does not lead to the equivalence of intensive and extensive margins of onsite (remote) productivity externality.[7] When cross-worksite spillover incurs no efficiency loss ($\tau = 1$), the intensive margin of onsite and remote productivity externalities are absent.

3.2.2 Sources of Externalities

The first-order condition for the socially optimal share of time spent working onsite (equation (11)) is different from the market equilibrium due to the existence of externalities. This section explains why these externalities exist and under what conditions they disappear.

Recall that in market equilibrium, the share of time spent working onsite is only determined by relative onsite productivity and relative onsite amenity ($\frac{\theta}{1-\theta} = \left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1}\left(\frac{\zeta}{\zeta^o}\right)^{\rho}$). Equation (11) shows three types of externalities that are not internalized in the market equilibrium: productivity externalities including onsite and remote productivity externalities (δ, δ^R), housing congestion ($\tilde{\gamma}$), and the composite of onsite and WFH disutility ($d(\theta)$).

For the productivity externalities, more onsite work or remote work lead to higher onsite or remote productivity. However, when workers allocate their time working onsite or WFH, they do not internalize that individual choices of labor deliver mode affect the aggregate productivity and take the onsite and remote productivity as constants. As a

[7] In the basic model, where all workers have the autonomy to determine whether to work onsite or WFH, $\tau = 0$ simplifies the onsite and remote labor shares to unity, thereby rendering the intensive and extensive margin elasticities identical. However, in a model with multi-mode workers, the condition $\tau = 0$ does not simplify the labor shares to 1, and thus the intensive and extensive margin of the productivity externalities may differ.

result, the share of time spent working onsite in market equilibrium deviates from the social optimum, resulting in a gap between market equilibrium and the social optimum in terms of efficiency labor units and total productivity. In this basic model with fixed total employment, when cross-worksite spillover incurs efficiency loss ($\tau \in [0, 1)$), the increase in share of time spent working onsite leads to a rise in the onsite productivity externality, along with decreases in the sharing of time spent WFH and the remote productivity externality. In other words, there is a trade-off between these two positive externalities because of total employment, working time constraints, and efficiency loss from onsite-remote collaboration. When the extensive margin of onsite and productivity externalities are absent ($\lambda = \lambda^R = 0$) or no efficiency loss in cross-worksite spillover ($\tau = 1$), the intensive margin of onsite and remote productivity externalities do not exist ($\delta = \delta^R = 0$).

Regarding housing congestion, workers take housing prices as given when making choices on the share of time spent working onsite. However, housing prices correlate with the share of time spent working onsite when the housing supply is not perfectly elastic ($\gamma \neq 1$). This correlation is because the share of time spent WFH affects the efficiency units of labor. An increasing efficiency unit of labor implies a rise in income and thus a higher demand for housing, which leads to increasing housing prices. The housing congestion is absent if the housing supply is perfectly elastic (when $\gamma = 1$, $\tilde{\gamma}$=1.) or the consumption share of housing is extremely low (when $1 - \alpha \rightarrow 0$, $\tilde{\gamma} \rightarrow 1$.).

The externality from the composite of disutility appears due to the optimum-market disparity created by productivity externalities and housing congestion. Specifically, in the presence of these two externalities, the social planner may choose a share of time spent working onsite that is different from the market equilibrium, which can result in a different composite of disutility in the social optimum compared to the market equilibrium. The externality from the composite of disutility is absent in two cases: (1) the productivity externalities and housing congestion are absent (When $\delta = \delta^R = 0$ and $\tilde{\gamma} = 1$, equation (11) becomes the same as the first-order condition in market equilibrium (equation 6)); (2) onsite and remote disutlities are equivalent (when $\zeta^o = \zeta$, $d(\theta) = 1$, which is irrelevant with θ). In the absence of three types of externalities, equation (11) becomes the same as equation (6), and market equilibrium coincides with the socially optimal allocation.

3.2.3 The Social Planner's Trade-offs

This section discusses the social planner's trade-offs. I start with the simplified case where there are only productivity externalities and then the case where three externalities are present. Finally, I derive the condition that the share of time spent working onsite in the social optimum larger, smaller, or coincides the market equilibrium in the presence of externalities.

If there are no externalities from housing congestion or the composite of onsite and WFH disutility, equation (11) simplifies to $\frac{\theta}{1-\theta} = \left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1}\left(\frac{\zeta+\delta}{\zeta^o+\delta^R}\right)^{\rho}$. The social planner chooses the socially optimal share of time spent working onsite to maximize aggregate productivity ($\ell(\theta)$), which is equivalent to maximizing utility in this case. When the onsite time elasticity is stronger than the remote time elasticity ($\delta > \delta^R$), the social planner maximizes the total productivity by choosing a higher onsite share than market equilibrium to utilize an increase in onsite productivity. Remote productivity will decrease, but the magnitude will be less than the increase in onsite productivity. When the intensive margin elasticity of onsite productivity equals the intensive margin elasticity of onsite productivity ($\delta = \delta^R$), the socially optimal share of time spent working onsite coincides with market equilibrium, which leads to the coincides of other optimal allocations and the market outcome[8]. This is a special case where the market equilibrium coincide with the social optimum when the strength of the two competing positive externalities are equal. This represents a new scenario where market equilibrium and the social optimum align, aside from cases where externalities are absent.

If all three externalities are present, the social planner balances the positive and negative effects of the externalities to choose an optimal share of time spent working onsite. For example, the social planner may choose a share of time spent working onsite that leads to higher productivity compared to market equilibrium at the costs of (1)

[8] When only productivity externalities exist, equating the simplified first-order condition for optimal onsite share to the one in market equilibrium yields $\frac{\zeta+\delta}{\zeta^o+\delta^R} = \frac{\zeta}{\zeta^o}$. There is no externality from disutility composite when onsite disutility is the same as WFH disutility ($\zeta^o = \zeta$); thus, the condition becomes $\frac{\zeta+\delta}{\zeta+\delta^R} = 1$. Simplifying this equation gives $\delta = \delta^R$.

increased housing prices and (2) increased composite of disutility if onsite disutility is larger than WFH amenity costs[9]. On the other hand, the social planner may choose a share of time spent working onsite that leads to lower productivity compared to market equilibrium with the gains from (1) decreased housing prices and (2) decreased composite of disutility if onsite disutility is smaller than WFH amenity costs. The gains in reduced disutility outweigh the decreased productivity cost, leading to higher social welfare.[10]

To compare the share of time spent working onsite in market equilibrium and socially optimal level in the presence of three externalities, I transform the first-order condition of optimal share of time spent working onsite as (Appendix C shows the derivation process):

$$\frac{\theta}{1-\theta} = \left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1} \left(\frac{\tilde{\gamma}(\delta+1)-\delta_d}{\tilde{\gamma}(\delta^R+1)-\delta_d^R}\right)^{\rho}, \tag{14}$$

where δ_d and δ_d^R are the elasticity of the disutility composite with respect to the share of time spent working onsite and share of time spent WFH, respectively:

$$\delta_d = \frac{\partial d(\theta)/d(\theta)}{\partial\theta/\theta} = \frac{(\zeta^o-\zeta)\theta}{d(\theta)}, \tag{15}$$

$$\delta_d^R = \frac{\partial d(\theta)/d(\theta)}{\partial(1-\theta)/(1-\theta)} = \frac{(\zeta-\zeta^o)(1-\theta)}{d(\theta)}. \tag{16}$$

In equation (14), $\tilde{\gamma}(\delta+1)-\delta_d$ incorporates the effect of changes in the share of time spent working onsite on housing price, onsite productivity, and the composite of disutility. $\tilde{\gamma}(\delta^R+1)-\delta_d^R$ incorporates the effect of changes in the share of time spent WFH on housing price, remote productivity, and the composite of disutility.

Comparing equation (14) with the first-order condition in the market equilibrium

[9] By the definition of the composite of onsite and WFH disutility $\frac{\partial d(\theta)}{\partial\theta} > 0$ if $\zeta^o > \zeta$.

[10] In the presence of disutility, the share of time spent working onsite chosen by the social planner does not necessarily result in higher productivity. To see this, assume the composite of onsite and WFH disutility does not vary by the share of time spent working onsite. Then, the first-order condition to determine the socially optimal share of time spent working onsite is $\frac{\theta}{1-\theta} = \left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1}\left(\frac{1+\delta}{1+\delta^R}\right)^{\rho}$. This equation is different from the condition when only productivity externalities are present: $\frac{\theta}{1-\theta} = \left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1}\left(\frac{\zeta+\delta}{\zeta^o+\delta^R}\right)^{\rho}$. The socially optimal share of time spent working onsite balances the gains and costs and thus may not increase productivity if the cost of increasing productivity is prohibitively expensive.

(equation (6)) gives the following conditions:

$$\theta^{soc} = \theta^{mkt} \Leftrightarrow \frac{\widetilde{\gamma}(\delta+1)-\delta_d}{\widetilde{\gamma}(\delta^R+1)-\delta_d^R} = \frac{\zeta}{\zeta^o}, \tag{17}$$

$$\theta^{soc} > \theta^{mkt} \Leftrightarrow \frac{\widetilde{\gamma}(\delta+1)-\delta_d}{\widetilde{\gamma}(\delta^R+1)-\delta_d^R} > \frac{\zeta}{\zeta^o}, \tag{18}$$

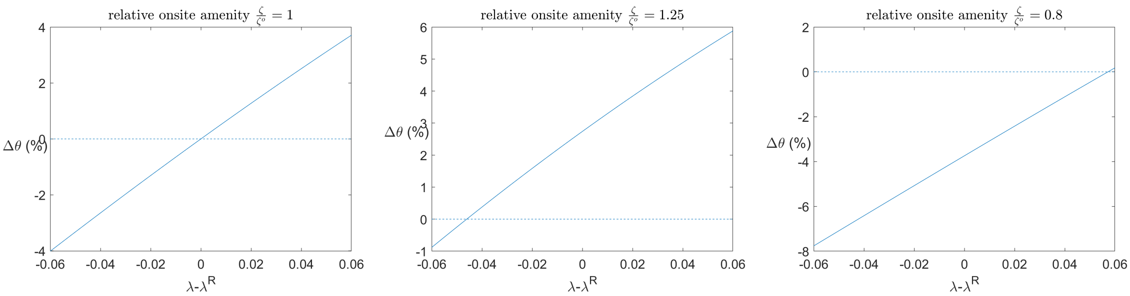
$$\theta^{soc} < \theta^{mkt} \Leftrightarrow \frac{\widetilde{\gamma}(\delta+1)-\delta_d}{\widetilde{\gamma}(\delta^R+1)-\delta_d^R} < \frac{\zeta}{\zeta^o}, \tag{19}$$

where θ^{soc} refers to the socially optimal share of time spent working onsite. θ^{mkt} refers to the share of time spent working onsite in the market equilibrium. Equation (17) gives the condition for the market outcome to coincide with the social optimum in the presence of externalities. Equation (18) implies that the social planner chooses a higher onsite share than market equilibrium when the adjusted intensive margin of the relative onsite productivity is larger than the relative onsite amenity (relative WFH disutility).

3.2.4 The Role of Productivity Externalities and Cross-worksite Spillover

This section explains how the extensive margins of onsite and remote productivity externalities and cross-worksite spillover affect the sign and magnitude of the gap between the social optimum and market equilibrium.

The gap of extensive margins of onsite and remote productivity externalities ($\lambda - \lambda^R$) affects the gap of intensive margins of onsite and remote productivity externalities ($\delta - \delta^R$) and, therefore, determines whether the socially optimal share of time spent working onsite is larger or smaller than market equilibrium. Figure (6) shows the relation between the optimum-market gap in the share of time spent working onsite and the gap between onsite and remote productivity externalities at the extensive margin in the presence of housing congestion and different scenarios of relative onsite amenity.



(a) (b) (c)

Figure 6: The Gap Between Onsite-Remote Productivity Externalities Affects the Direction of the Optimum-Market Gap

Note: The gap in the share of time spent working onsite between the social optimum and the market outcome $\Delta\theta = \frac{\theta^{soc}}{\theta^{mkt}} - 1 \times (100\%)$. Parameter values: $\gamma = 0.09, \tau = 0.015, \lambda = 0.06, \alpha = 0.76, \rho = 1.37, \bar{b} = \bar{a} = 1$. In this case, the cross-worksite spillover is small, so the gap between onsite and remote productivity externalities at the intensive and extensive margin are very close ($\lambda - \lambda^R \approx \delta - \delta^R$).

In figure (6(a)), onsite disutility and WFH disutility are the same, resulting in the composite of disutility being a constant and absence of externality from the composite of disutility. In this case, when extensive margins of onsite and remote productivity externalities have the same strength ($\lambda = \lambda^R$), the market equilibrium coincides with the social optimum. In figure (6(b)), onsite disutility is smaller than WFH disutility. Therefore, an increase in the share of time spent working onsite means a decrease in the composite of onsite and WFH disutility. Suppose the social planner increases the share of time spent working onsite; workers benefit from a reduction of the composite of onsite and WFH disutility, which leads to an increase in the left-hand side of the equation (17). The onsite productivity externality should be smaller than the remote productivity externality to keep the equation balanced. The intuition is that when the cost of increasing the share of time spent working onsite is relatively low, the social planner will choose more onsite work even if the onsite productivity externality is relatively weaker. Figure (6(c)) shows that when onsite disutility is higher than WFH disutility, the strength of onsite productivity externality needs to be strong enough for the social planner to choose a share of time spent working onsite higher than market equilibrium. In other words, if the cost of more onsite work is relatively high, the benefit from onsite productivity spillover needs to be strong enough for the social planner to choose a share of time spent working onsite larger than the market equilibrium.

The gap between the extensive margins of onsite and remote productivity externalities $(\lambda - \lambda^R)$ also affects the magnitude of the gap between the market outcome and the social optimum. Figure 7(a) shows a situation without the externality from the composite of disutility. The blue line describes how the gap of share of time spent working onsite between the social optimum and market equilibrium changes as $\lambda - \lambda^R$ changes. The larger the gap in extensive margins of onsite and remote productivity externalities, the larger the gap in θ^{soc} and θ^{mkt}. As shown in the orange lines, the gap in the share of time spent working onsite leads to gaps in productivity, income, and utility between market equilibrium and the social optimum.

As the gap between the social optimum and market equilibrium level of onsite time widens, the required subsidy to achieve the social optimum increases. Consider two types of policy: The first is implementing income tax to subsidies onsite work or remote work. The disposable income under this policy is

$$\mathscr{W}(\theta) = I(\theta)h(\theta), \quad h(\theta) = (1 - \mathscr{T})\,(1 + s\theta)\,, \tag{20}$$

where $\mathscr{T}$ is income tax rate. s is the onsite subsidy (tax) rate. $s > 0$ corresponds to subsidizing onsite work or taxing remote work; $s < 0$ corresponds to taxing onsite work or subsidizing remote work.

The second is levying externality tax to fund the subsidy. For example, collect remote tax (income tax adjusted by the share of time spent WFH) and use the tax to subsidize onsite work. If the social planner taxes (subsidies) remote work and subsidies (taxes) the onsite work, the disposable income becomes

$$\mathscr{W}(\theta) = I(\theta)g(\theta), \quad g(\theta) = (1 - \mathscr{t}(1 - \theta))\,(1 + t\theta)\,, \tag{21}$$

where $\mathscr{t}$ is the tax (subsidy) rate based on share of time spent WFH, t is subsidy (tax) rate based on share of time spent working onsite. (See Appendix (D) for derivation details.)

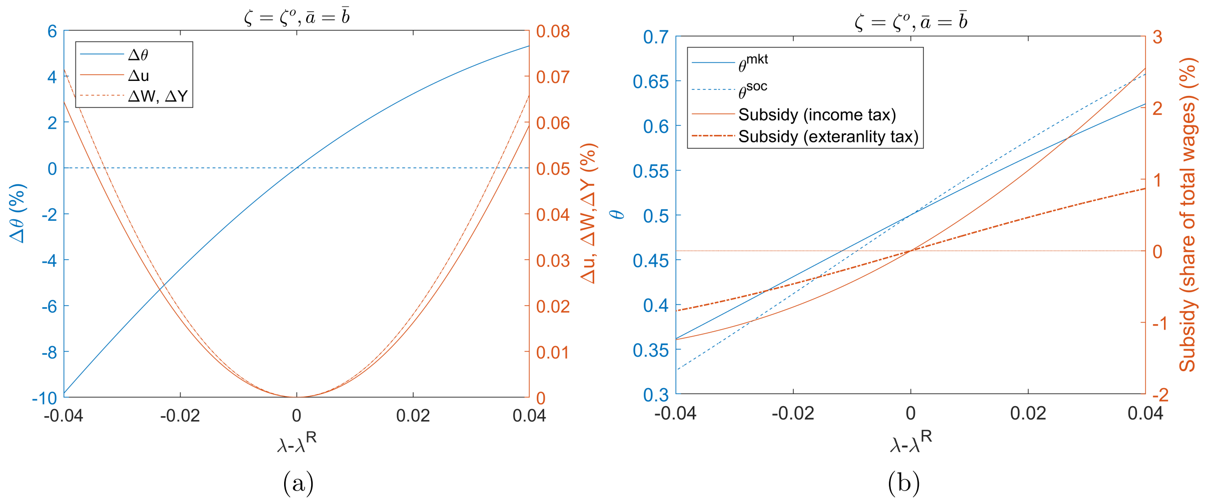
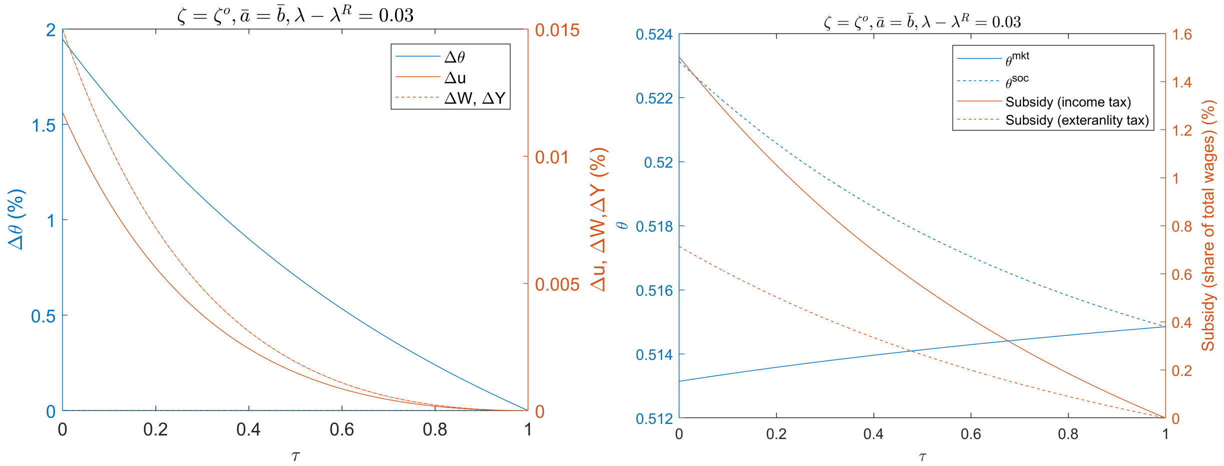


Figure 7: The Gap Between Onsite-Remote Productivity Externalities Affects the Magnitude of the Optimum-Market Gap
Note: $\Delta x = (\frac{x^{soc}}{x^{mkt}} - 1) \times 100\%$. Parameter values: γ=0.09, τ=0.015, λ=0.06, α=0.76, ρ=1.37, $\bar{b} = \bar{a} = 1$.

Figure 7(b) shows that the gap in the extensive margins of onsite-remote productivity externalities causes the share of time spent working onsite in the market equilibrium (solid blue line) to deviate from the social optimum (dotted blue line), leading to a larger optimal subsidy (orange lines).

The cross-worksite spillover (τ) affects the magnitude of the optimum-market gap and the subsidy. If there is no efficiency loss in cross-worksite spillover, the cross-worksite spillover becomes one. As a result, onsite and remote productivity externalities are irrelevant to how workers allocate work time onsite or at home, implying the absence of intensive margins of productivity externalities. Figure (8(a)) shows the relation between the gap between the social optimum and the market outcome in a scenario where the externality from the composite of disutility is absent. As the cross-worksite spillover increases, the gap become smaller. Figure (8(b)) shows that as the cross-worksite spillover increases, the gap between market equilibrium and the social optimum level of onsite time and the socially optimal subsidy decrease.



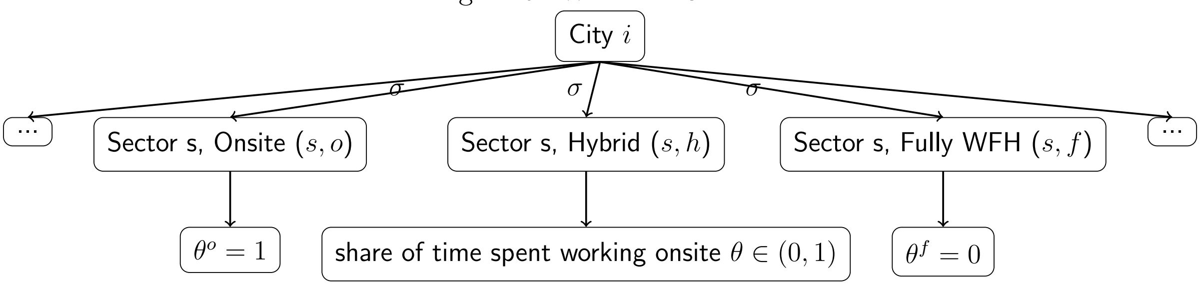
(a) (b)

Figure 8: Cross-worksite Spillover Affects the Magnitude of the Optimum-Market Gap
Note: $\Delta x = (\frac{x^{soc}}{x^{mkt}} - 1) \times 100\%$. Parameter values: $\gamma = 0.09$, τ=0.015, α=0.76, ρ=1.37, $\bar{b} = \bar{a} = 1$, λ=0.06, λ^R=0.03.

4 The Full Model

4.1 The Market Equilibrium

Figure 9: Workers' Choices



Note: A worker living and working in city i choose a sector s and one of three work modes. If the worker chooses working onsite or fully WFH, her share of time spent working onsite is determined as 1 or 0. If the worker chooses hybrid, she then chooses a share of time spent working onsite $\theta \in (0, 1)$. σ measures the dispersion of preferences for sectors-work mode pairs.

The economy has $i = 1, .., N$ cities. Each city has a fixed size of employment ($\bar{L}_i$). Workers live where they work. There is no migration across cities, but there is mobility across sectors $s \in \{1, 2, ..., S\}$ and work modes. Each worker chooses a pair of sector-work mode (s, m), where the work modes are ($m \in \{o, h, f\}$) fully onsite (o), hybrid(h), and fully WFH (f). If workers choose hybrid work, they then decide the share of time spent working onsite, $\theta_{is} \in (0, 1)$, and the share of time spent WFH, $1 - \theta_{is}$.[11]

[11] Onsite workers' share of time spent working onsite $\theta^o_{is} = 1$. Fully WFH workers' share of time spent

Each city produces a tradable good that uses labor inputs from all sectors in the city. Tradable goods produced by cities are freely traded, and they are perfect substitutes. Workers consume tradable goods and housing. Housing is produced using parts of tradable goods and land as inputs. The returns to land is allocated to local workers. At the city-sector level, productivity spillover takes place both within and between onsite and remote labor, and onsite-remote knowledge-sharing may incur an efficiency loss.

4.1.1 Preference

The utility for a worker ω who lives and works in city i and works in sector s, and choose an work mode m is:

$$u_{is\omega}^{m} = \frac{(c_{is}^{m})^{\alpha}(\hbar_{is}^{m})^{1-\alpha}\epsilon_{is}^{m}}{\bar{\phi} d_{is}^{m}(\theta_{is}^{m})} z_{(s,m)\omega}, \tag{22}$$

where c_{is}^{m} and $\hbar_{is}^{m}$ are consumption for tradable goods and housing, respectively. $\bar{\phi} = \alpha^{\alpha}(1-\alpha)^{1-\alpha}$ is a constant. The disutility $d_{is}^{m}(\theta_{is}^{m})$ varies by city, sector, and work mode. For onsite workers, disutility is the commuting cost; for hybrid workers, disutility combines the commuting cost and WFH amenity cost; for workers fully WFH, their disutility is the WFH amenity cost:

$$d_{is}^{m}(\theta_{is}^{m}) = \begin{cases} e^{kt_i} & \text{if } m = o, \\ e^{kt_i}\theta_{is} + \zeta_{is}(1-\theta_{is}) & \text{if } m = h, \\ \zeta_{is} & \text{if } m = f. \end{cases} \tag{23}$$

ζ_{is} denotes the WFH amenity cost. e^{kt_i} refers to the commuting cost, where t_i is the city-specific commuting time and k denotes the elasticity of commuting cost to commuting time.

ϵ_{is}^{m} is exogenous shifters that govern workers' preference for sector-work mode $(s-m)$ choices. $z_{(s,m)\omega}$ is the worker's idiosyncratic preference for a sector-work mode pair. Assume $z_{(s,m)\omega}$ is independent for all individuals and follows the Fréchet distribution with cumulative distribution function $F_{z_{(s,m)\omega}}(z) = e^{-z^{-\sigma}}$. σ governs the dispersion of

working onsite $\theta_{is}^{f} = 0$. For hybrid workers, $\theta_{is}^{h} \in (0,1)$. For simplicity, I ignore the subscript h and use θ_{is} to refer to share of time hybrid workers spend working onsite.

preference for sector-work mode pairs.

By utility maximization, the indirect utility is:

$$u_{is\omega}^{m} = \frac{I_{is}^{m}(\theta_{is}^{m})\epsilon_{is}^{m}}{d_{is}^{m}(\theta_{is}^{m})P^{\alpha}q_{i}^{1-\alpha}} z_{(s,m)\omega}, \tag{24}$$

where the worker's income $I_{is}^{m}(\theta_{is}^{m})$ consists of wages and returns to land. W_{is}^{m} denote wages and R_{is}^{m} denote returns to land. Assume the returns to land is distributed locally and is proportion to a worker's total wages: $R_{is}^{m} = \bar{r}W_{is}^{m}$. This assumption ensures that worker's choices and labor supply are not affected by the allocation of returns to land. The ratio of returns to land to total wages, $\bar{r}$, is a constant related to the share of tradable goods in consumption (α) and housing supply elasticity. To summarize, the income is $I_{is}^{m} = (1+\bar{r})W_{is}^{m}$. P is the price of tradable goods, which is equalized across cities due to free trade. The price of tradable goods is normalized to 1. q_i is the housing price in city i. Wages for workers are multiplications of city-sector-specific unit wages (w_{is}) and work mode-specific efficiency labor unit:

$$W_{is}^{m}(\theta_{is}^{m}) = \begin{cases} w_{is}B_{is} & \text{if } m = o, \\ w_{is}\beta_{is}\ell_{is}(\theta_{is}) & \text{if } m = h, \\ w_{is}A_{is} & \text{if } m = f. \end{cases} \tag{25}$$

Hybrid worker's efficiency labor unit $\ell_{is}(\theta_{is})$ combines the onsite and remote productivity:

$$\ell_{is}(\theta_{is}) = [(A_{is}(1-\theta_{is}))^{\frac{\rho-1}{\rho}} + (B_{is}\theta_{is})^{\frac{\rho-1}{\rho}}]^{\frac{\rho}{\rho-1}}, \tag{26}$$

where $\rho > 1$ is the elasticity of substitution between working onsite and WFH. A_{is} refers to remote productivity and B_{is} denotes onsite productivity. Workers consider onsite and remote productivity to be constant. β_{is} refers to the exogenous hybrid productivity that cannot be explained by w_{is} and $\ell_{is}(\theta_{is})$.

Hybrid workers choose the share of time onsite (θ_{is}) by utility maximization. The share of time spent working onsite depends on relative productivity and relative amenity of onsite work:

$$\frac{\theta_{is}}{1-\theta_{is}} = \left(\frac{B_{is}}{A_{is}}\right)^{\rho-1} \left(\frac{\zeta_{is}}{e^{kt_i}}\right)^{\rho}. \tag{27}$$

Rearranging this equation gives the solution for share of time hybrid workers spend working onsite in the market equilibrium:

$$\theta_{is} = \frac{1}{1 + \left(\frac{A_{is}}{B_{is}}\right)^{\rho-1} \left(\frac{e^{kt_i}}{\zeta_{is}}\right)^{\rho}}. \tag{28}$$

4.1.2 Labor Supply

Since the idiosyncratic preference for sector-work mode follows Fréchet distribution, the labor supply for workers in city i, sector s with work mode m is:

$$L_{is}^m = \left(\frac{\phi_{is}^m}{\Phi_i}\right)^{\sigma} \bar{L}_i, m \in \{o, h, f\}, \tag{29}$$

where $\phi_{is}^m = \frac{I_{is}^m \epsilon_{is}^m}{d_{is}^m(\theta_{is}^m)}$, $\Phi_i = [\sum_s \sum_m (\phi_{is}^m)^{\sigma}]^{\frac{1}{\sigma}}$. The employment in city i and sector s can be derived as

$$L_{is} = \sum_s L_{is}^m = \frac{1}{\Phi_i^{\sigma}} \sum_m \left(\frac{I_{is}^m \epsilon_{is}^m}{d_{is}^m(\theta_{is}^m)}\right)^{\sigma} \bar{L}_i, \tag{30}$$

Define $\widetilde{\epsilon}_{is}^m = \frac{\epsilon_{is}^m}{\sum_m \epsilon_{is}^m}$, then city-sector employment can be expressed as

$$L_{is} = \frac{\epsilon_{is}}{\Phi_i^{\sigma}} \sum_m \left(\frac{I_{is}^m \widetilde{\epsilon}_{is}^m}{d_{is}^m(\theta_{is}^m)}\right)^{\sigma} \bar{L}_i, \tag{31}$$

where $\epsilon_{is} = (\sum_m \epsilon_{is}^m)^{\sigma}$ is the city-sector specific labor supply shifter. This shifter will be used as an instrumental variable for city-sector employment in the structural estimation in section 5. Combining equation (29) and (30), labor supply for workers in city i, sector s with work mode m can be expressed as a function of city-sector employment:

$$L_{is}^m = \left(\frac{\phi_{is}^m}{\Phi_{is}}\right)^{\sigma} L_{is}, m \in \{o, h, f\}, \tag{32}$$

where $\Phi_{is} = [\sum_m (\phi_{is}^m)^{\sigma}]^{\frac{1}{\sigma}}$.

4.1.3 Production

The production function of the tradable good in city i combines outputs from all sectors:

$$Y_i = \left[\sum_s \left(\bar{y}_{is} (y_{is})^{\frac{\eta-1}{\eta}} \right) \right]^{\frac{\eta}{\eta-1}}, \tag{33}$$

where $\bar{y}_{is}$ is the city-sector specific productivity shifters. Assume constant returns to scale, thereby $\sum_s \bar{y}_{is} = 1$. η is the elasticity of substitution between sectors. The output in city i sector s (y_{is}) is a linear combination of outputs from workers in different work modes:

$$y_{is} = y_{is}^o + y_{is}^h + y_{is}^f = B_{is} L_{is}^o + \beta_{is} \ell_{is}(\theta_{is}) L_{is}^h + A_{is} L_{is}^f. \tag{34}$$

By profit maximization, the city-sector-specific unit wage is:

$$w_{is} = P \bar{y}_{is} \left(\frac{Y_i}{y_{is}} \right)^{\frac{1}{\eta}}. \tag{35}$$

Onsite, hybrid, and fully WFH workers receive the following total wages, respectively:

$$W_{is}^o = w_{is} B_{is}, \quad W_{is}^h = w_{is} \beta_{is} \ell_{is}(\theta_{is}), \quad W_{is}^f = w_{is} A_{is}. \tag{36}$$

The production function of housing in city i is:

$$H_i = K_i^{\gamma} \bar{H}_i^{1-\gamma}, \tag{37}$$

where K_i is the tradable goods used as an input to produce housing. $\bar{H}_i$ refers to land. γ governs the elasticity of the housing supply. By profit maximization, the demand for tradable goods used to produce housing is

$$K_i = \gamma \frac{q_i}{P} H_i, \tag{38}$$

where q_i is the housing price. The demand for land $\bar{H}_i = \frac{1-\gamma}{\gamma} \frac{P}{r_i} K_i$ implies the total returns to land in city i is

$$R_i = r_i \bar{H}_i = \frac{1-\gamma}{\gamma} P K_i, \tag{39}$$

where r_i is the per unit returns to land.

4.1.4 Externalities

When making decisions, workers do not account for the impact of their choices on aggregate productivity, which in turn affects other workers. When firms in different sectors choose labor to maximize their profits, they do not internalize the productivity externalities among workers. In other words, workers and firms treat the onsite productivity B_{is} and remote productivity A_{is} as constant when making choices.

Before introducing the function forms of onsite and remote productivity, I define the aggregation of onsite and remote labor. They are both contributed by two types of workers. The total labor for onsite work (L^B_{is}) in the city i, sector s combines labor from onsite workers and hybrid workers when they work onsite:

$$L^B_{is}(\theta_{is}) = L^o_{is} + \theta_{is} L^h_{is}. \tag{40}$$

The remote labor (L^A_{is}) combines labor from fully WFH workers and hybrid workers when they WFH:

$$L^A_{is}(\theta_{is}) = L^f_{is} + (1-\theta_{is}) L^h_{is}. \tag{41}$$

The productivity for onsite work is:

$$B_{is} = \bar{b}_{is} \left[L^B_{is}(\theta_{is}) + \tau L^A_{is}(\theta_{is}) \right]^{\lambda}. \tag{42}$$

The productivity for remote work is:

$$A_{is} = \bar{a}_{is} [L^A_{is}(\theta_{is}) + \tau L^B_{is}(\theta_{is})]^{\lambda^R}. \tag{43}$$

where $\bar{b}_{is}$ and $\bar{a}_{is}$ are exogenous onsite and remote productivity, respectively. λ and λ^R are the elasticities of onsite and remote productivity to aggregate labor. They govern the strength of onsite and remote productivity externalities at the extensive margin, respectively.

$\tau \in [0, 1]$ refers to cross-worksite spillover. As in the simple model, when $\tau = 0$, onsite and remote labor contribute separately to onsite and remote productivity. When $\tau = 1$, onsite and remote productivity are not directly related to the choice of work mode or to share of time hybrid workers spend working onsite ($B_{is} = \bar{b}_{is} L_{is}^{\lambda}$, $A_{is} = \bar{a}_{is} L_{is}^{\lambda^R}$). The aggregate productivity is affected by workers' choices in work modes and share of time spent working onsite when these choices lead to a labor relocation across sectors and change the employment of a sector (L_{is}). $\tau \in (0, 1)$ implies an efficiency loss in productivity spillover between onsite and remote workers.

4.1.5 Market Clearing

The labor market clearing condition for workers is:

$$L_{is}^m = \left(\frac{\phi_{is}^m}{\Phi_i} \right)^{\sigma} \bar{L}_i, m \in \{o, h, f\} \tag{44}$$

where $\phi_{is}^m = \frac{I_{is}^m \epsilon_{is}^m}{d_{is}^m(\theta_{is}^m)}$, $\Phi_i = [\sum_s \sum_m (\epsilon_{is}^m)^{\sigma}]^{\frac{1}{\sigma}}$. $I_{is}^m = (1 + \bar{r}) W_{is}^m$. $W_{is}^o = w_{is} B_{is}, W_{is}^h = w_{is} \beta_{is} \ell_{is}(\theta_{is}), W_{is}^f = w_{is} A_{is}$. The unit wage $w_{is} = P \bar{y}_{is} \left(\frac{Y_i}{y_{is}} \right)^{\frac{1}{\eta}}$.

The housing market clearing condition is:

$$\sum_s \sum_{m \in \{o,h,f\}} \hbar_{is}^m L_{is}^m = H_i \tag{45}$$

$$\frac{1 - \alpha}{q_i} I_i = H_i, \tag{46}$$

where the total income in city i combines the total wage and returns to land $I_i = W_i + R_i$. The total wages are $W_i = \sum_s \sum_m (W_{is}^m L_{is}^m)$. Combining equation (46) and (38), the demand for tradable goods used to produce housing can be expressed as a function of total income:

$$K_i = \gamma(1 - \alpha) \frac{I_i}{P}. \tag{47}$$

Substituting this equation and the housing production function (37) to equation (46), the housing price can be solved as:

$$q_i = \left(\frac{P}{\gamma}\right)^{\gamma} \left(\frac{(1-\alpha)I_i}{\bar{H}_i}\right)^{1-\gamma}. \tag{48}$$

Substituting equation (47) to equation (39), the total returns to land in city i can be expressed as $R_i = (1-\gamma)(1-\alpha)I_i = (1-\gamma)(1-\alpha)(W_i + R_i)$. Rearranging the equation yields a relation between returns to land and total wages: $R_i = \frac{(1-\gamma)(1-\alpha)}{1-(1-\gamma)(1-\alpha)}W_i$. By assumption, returns to land allocated to all workers has the same proportion r relative to their wages, $R_i = rW_i$. Therefore, the proportion of land returns to wages is:

$$\bar{r} = \frac{(1-\gamma)(1-\alpha)}{1-(1-\gamma)(1-\alpha)} = \frac{1}{\alpha+\gamma(1-\alpha)} - 1. \tag{49}$$

A worker's income is $I_{is}^m = (1+\bar{r})W_{is}^m = \frac{1}{\alpha+\gamma(1-\alpha)}W_{is}^m$. Total income in city i is $I_i = \frac{W_i}{\alpha+\gamma(1-\alpha)}$. Substituting $I_i = \frac{W_i}{\alpha+\gamma(1-\alpha)}$ and zero profit condition ($W_i = PY_i$) into $K_i = \gamma(1-\alpha)\frac{I_i}{P}$, the tradable goods used to produce housing can be expressed as a function of total outputs $K_i = \frac{\gamma(1-\alpha)}{\gamma(1-\alpha)+\alpha}Y_i$.

The goods market clearing condition is:

$$P\sum_s \sum_{m\in\{o,h,f\}} c_{is}^m L_{is}^m + PK_i = PY_i \tag{50}$$

$$\alpha\frac{I_i}{P} + K_i = Y_i \tag{51}$$

By substituting $K_i = \gamma(1-\alpha)\frac{I_i}{P}$ and $I_i = \frac{W_i}{\alpha+\gamma(1-\alpha)}$ to equation (51) gives $W_i = PY_i$. This condition is satisfied according to the zero profit condition.

4.1.6 The Market Equilibrium

Given parameters $\{k, \rho, \sigma, \eta, \gamma, \lambda^R, \lambda, \tau\}$, exogenous variables $\{t_i, \bar{L}_i, \bar{H}_i\}$, and exogenous shifters $\{\zeta_{is}, \epsilon_{is}^o, \epsilon_{is}^h, \epsilon_{is}^f, \beta_{is}, \bar{y}_{is}, \bar{a}_{is}, \bar{b}_{is}\}$, a competitive allocation consists of the share of time spent working onsite chosen by hybrid workers (θ_{is}), employment ($L_{is}^o, L_{is}^h, L_{is}^f$), total wages ($W_{is}^o, W_{is}^h, W_{is}^f$), tradable goods used to produce housing (K_i), local returns to land (R_i), housing price (q_i), onsite productivity (B_{is}) and remote productivity (A_{is}) such that equations (26), (28), (29), (36), (38), (39), (42), (43), (45), and (48) are satisfied.

4.2 Welfare

The expected utility for a worker in city i is:

$$U_i = \Phi_i \frac{\Gamma(1-\frac{1}{\sigma})}{P^{\alpha} q_i^{1-\alpha}}, \tag{52}$$

where $\Phi_i = [\sum_s \sum_m (\phi_{is}^m)^{\sigma}]^{\frac{1}{\sigma}}$, $\phi_{is}^m = \frac{I_{is}^m \epsilon_{is}^m}{d_{is}^m(\theta_{is}^m)}$. $I_{is}^m = \frac{W_{is}^m}{\alpha+\gamma(1-\alpha)}$. $W_{is}^o = w_{is}B_{is}$, $W_{is}^h = w_{is}\beta_{is}\ell_{is}(\theta_{is})$, $W_{is}^f = w_{is}A_{is}$.

The social welfare for the economy is defined as:

$$U = \sum_i U_i \bar{L}_i. \tag{53}$$

4.3 The Socially Optimal Labor Allocations

The social planner internalizes externalities and solves for allocations in the share of time spent working onsite for hybrid workers, employment, consumption, housing, tradable goods used to produce housing, and trade flow between cities to maximize the total welfare of all cities (See Appendix E for derivation details). The first-order conditions determining the socially optimal allocation differ from those of the market equilibrium in terms of both the intensive and extensive margins of labor.

4.3.1 The Optimal Share of Time Spent Working Onsite For Hybrid Workers

The equation that characterizes the socially optimal share of time spent working onsite for hybrid workers is:

$$\frac{\theta_{is}^*}{1-\theta_{is}^*} = \left(\frac{B(\theta_{is}^*)}{A(\theta_{is}^*)}\right)^{\rho-1} \left[\frac{\zeta_{is} + d_{is}(\theta_{is}^*)\left(F_{is}\left(\overbrace{\delta_{is}}^{\text{onsite time elasticity}} +1\right)-1\right)}{e^{kt_i} + \underbrace{d_{is}(\theta_{is}^*)}_{\substack{\text{hybid worker's}\\ \text{composite of}\\ \text{disutility}}}\left(\underbrace{F_{is}}_{\substack{\text{housing congestion,}\\ \text{onsite-remote/sector}\\ \text{productivity spillover}}}\left(\underbrace{\delta_{is}^R}_{\substack{\text{intensive margin of}\\ \text{remote productivity}\\ \text{externality}}} +1\right)-1\right)}\right]^{\rho}, \tag{54}$$

where

$$\delta_{is} = \frac{\partial B_{is}(\theta_{is})/B_{is}}{\partial \theta_{is}(\theta_{is})/\theta_{is}} = \lambda(1-\tau)\frac{\theta_{is}L_{is}^{h}}{L_{is}^{B}+\tau L_{is}^{A}}, \tag{55}$$

$L_{is}^{B}(\theta_{is}) = L_{is}^{o}+\theta_{is}L_{is}^{h}, L_{is}^{A}(\theta_{is}) = L_{is}^{f}+(1-\theta_{is})L_{is}^{h}$. δ_{is} is the elasticity of onsite productivity at city i sector s with respect to share of time hybrid workers spend working onsite, which governs the onsite time elasticity that came from hybrid workers.

$$\delta_{is}^{R} = \frac{\partial A_{is}(\theta_{is})/A_{is}}{\partial(1-\theta_{is})/(1-\theta_{is})} = \lambda^{R}(1-\tau)\frac{(1-\theta_{is})L_{is}^{h}}{L_{is}^{A}+\tau L_{is}^{B}}. \tag{56}$$

δ_{is}^{R} is the elasticity of remote productivity at city i sector s with respect to hybrid workers' share of time spent WFH, which governs the intensive margin of remote externality that came from hybrid workers. The onsite (remote) time elasticity is calculated by multiplying three factors: the extensive margin of onsite (remote) productivity externality (λ, λ^{R}), the cross-worksite efficiency loss $(1-\tau)$, and the ratio of hybrid workers' onsite (remote) labor $(\theta_{is}L_{is}^{h}, (1-\theta_{is})L_{is}^{h})$ to the aggregate labor that contributes to onsite (remote) productivity $(L_{is}^{B}+\tau L_{is}^{A}$, $L_{is}^{A}+\tau L_{is}^{B})$. The intensive margin of productivity externalities varies by city and sector because of the variations in labor composition and share of time hybrid workers spend working onsite.

Note that a stronger onsite agglomeration elasticity relative to the remote agglomeration elasticity $(\lambda > \lambda^{R})$ does not necessarily guarantee that the onsite time elasticity is larger than the remote time elasticity $(\delta > \delta^{R})$. For example, when a city and sector has a large share of onsite workers and low shares of hybrid and remote workers, it is possible that $\lambda > \lambda^{R}$ but $\delta < \delta^{R}$ due to the share of hybrid workers' onsite labor is smaller than the share of hybrid workers' remote labor $(\frac{\theta_{is}L_{is}^{h}}{L_{is}^{B}+\tau L_{is}^{A}} < \frac{(1-\theta_{is})L_{is}^{h}}{L_{is}^{A}+\tau L_{is}^{B}})$. The intuition is that a high number of onsite workers already creates a strong onsite productivity effect, such that an increase in the share of time spent working onsite by hybrid workers only marginally increases onsite productivity. However, if hybrid workers allocate more time WFH, it could result in a relatively larger boost to remote productivity.

$$F_{is} = \frac{\widetilde{P}\frac{\partial Y_i}{\partial \theta_{is}}}{\widetilde{I}_{is}^{h}L_{is}^{h}\frac{\partial \ell_{is}(\theta_{is})/\partial \theta_{is}}{\ell_{is}(\theta_{is})}}, \widetilde{P} = \frac{(\alpha+\gamma(1-\alpha))\sum_i\sum_s\sum_m \widetilde{I}_{is}^{m}L_{is}^{m}}{\sum_i Y_i} \tag{57}$$

F_{is} refers to the onsite-remote and cross-sector productivity spillover effect driven by the change of share of time hybrid workers spend working onsite in sector s. $\widetilde{I}_{is}^m$ denotes the total expenditure in the solution to the social planner's problem. The numerator $\widetilde{P}\frac{\partial Y_i}{\partial \theta_{is}}$ measures the change in the value of the tradable goods in city i when hybrid workers change share of time spent working onsite. It reflects the marginal total product value of hybrid workers' onsite work. The denominator $\widetilde{I}_{is}^h L_{is}^h \frac{\partial \ell_{is}(\theta_{is})/\partial \theta_{is}}{\ell_{is}(\theta_{is})}$ measures the change in total income of the hybrid workers when they adjust share of time spent working onsite. It represents the value of hybrid workers' marginal product from onsite work. Thus, the ratio of the two terms reflects how much the value of tradable goods changes, excluding the change in the value of the hybrid worker's product.

To illustrate the productivity spillover effect, consider the following example: An increase in share of time hybrid workers spend working onsite in sector s increases onsite productivity but decreases remote productivity of sector s, resulting in a change in the total productivity of sector s. If the total productivity of this sector increases, it causes a relative decline in the productivity of other sectors. In the case where there is only one city and one sector and all workers are hybrid, F_{is} simplifies to $\alpha + \gamma(1 - \alpha)$, which only includes the housing congestion effect, as in the basic model.

Similar to the basic model (3.2.3), when determining the socially optimal share of time spent working onsite, the social planner balances the effect on the hybrid worker's composite of disutility, housing congestion, as well as intensive margins of onsite and remote productivity externalities. In addition, the social planner also considers the productivity spillover across work modes and sectors.

To understand how the onsite-remote and sector productivity spillover affect the social planner's trade-offs, consider a simplified scenario in which only this externality and productivity externalities exist. In this case, the social planner may choose a share of time spent working onsite in a sector even if it results in lower productivity for hybrid workers in this sector, as long as the benefits to other work modes or sectors outweigh the loss. This mechanism operates through the labor reallocation channel. For example, suppose a change in the share of time spent working onsite reduces the productivity of hybrid workers in sector s, causing hybrid workers' wages in sector s to decrease. As a consequence, workers reallocate to other sector-work mode cells, thereby increasing productivity in those sectors and work modes. If the productivity increases in other sector-work mode cells exceed the hybrid workers' productivity decrease in sector s, then the social planner will implement this change in the share of time spent working onsite.

Equation (54) can be expressed as

$$\frac{\theta_{is}}{1-\theta_{is}} = \left(\frac{B_{is}(\theta_{is})}{A_{is}(\theta_{is})}\right)^{\rho-1} \left(\frac{F_{is}(\delta_{is}+1)-\delta_{d,is}}{F_{is}(\delta_{is}^{R}+1)-\delta_{d,is}^{R}}\right)^{\rho}, \tag{58}$$

where $\delta_{d,is}$ and $\delta_{d,is}^{R}$ are the elasticity of the hybrid worker's composite of disutility with respect to their share of time spent working onsite and share of time spent WFH in city i sector s, respectively:

$$\delta_{d,is} = \frac{\partial d_{is}(\theta_{is})/d_{is}(\theta_{is})}{\partial\theta_{is}/\theta_{is}} = \frac{(e^{kt_i}-\zeta_{is})\theta_{is}}{d_{is}(\theta_{is})}, \tag{59}$$

$$\delta_{d,is}^{R} = \frac{\partial d_{is}(\theta_{is})/d_{is}(\theta_{is})}{\partial(1-\theta_{is})/(1-\theta_{is})} = \frac{(\zeta_{is}-e^{kt_i})(1-\theta_{is})}{d_{is}(\theta_{is})}. \tag{60}$$

$\frac{F_{is}(\delta_{is}+1)-\delta_{d,is}}{F_{is}(\delta_{is}^{R}+1)-\delta_{d,is}^{R}}$ represents the relative onsite time elasticity adjusted by the other sources of externalities.

Comparing equation (58) with the first-order condition in market equilibrium (equation (27)) gives the following conditions:

$$\theta_{is}^{soc} > \theta_{is}^{mkt} \Leftrightarrow \left(\frac{B_{is}(\theta_{is}, L_{is,soc}^{m})}{A_{is}(\theta_{is}, L_{is,soc}^{m})}\right)^{\rho-1} \left(\frac{F_{is}(\delta_{is}+1)-\delta_{d,is}}{F_{is}(\delta_{is}^{R}+1)-\delta_{d,is}^{R}}\right)^{\rho} > \left(\frac{B_{is}(\theta_{is}, L_{is,mkt}^{m})}{A_{is}(\theta_{is}, L_{is,mkt}^{m})}\right)^{\rho-1} \left(\frac{\zeta_{is}}{e^{kt_i}}\right)^{\rho}, \tag{61}$$

$$\theta_{is}^{soc} < \theta_{is}^{mkt} \Leftrightarrow \left(\frac{B_{is}(\theta_{is}, L_{is,soc}^{m})}{A_{is}(\theta_{is}, L_{is,soc}^{m})}\right)^{\rho-1} \left(\frac{F_{is}(\delta_{is}+1)-\delta_{d,is}}{F_{is}(\delta_{is}^{R}+1)-\delta_{d,is}^{R}}\right)^{\rho} < \left(\frac{B_{is}(\theta_{is}, L_{is,mkt}^{m})}{A_{is}(\theta_{is}, L_{is,mkt}^{m})}\right)^{\rho-1} \left(\frac{\zeta_{is}}{e^{kt_i}}\right)^{\rho}. \tag{62}$$

where θ_{is}^{soc} and θ_{is}^{mkt} denotes the share of time hybrid workers spend working onsite for hybrid workers in the social optimum and market equilibrium, respectively. $L_{is,soc}^{m}$ and

$L^m_{is,mkt}$ denote employment in the social optimum and market equilibrium, respectively. Conditions (61) and (62) suggest that whether optimal onsite share larger than market equilibrium depend on two aspects: (1) whether adjusted relative onsite time elasticity is larger than relative onsite amenity ($\frac{F_{is}(\delta_{is}+1)-\delta_{d,is}}{F_{is}(\delta^R_{is}+1)-\delta^R_{d,is}}$ v.s. $\frac{\zeta_{is}}{e^{kt_i}}$). (2) whether relative onsite productivity in the social optimum is larger than market equilibrium ($\frac{B_{is}(\theta_{is},L^m_{is,soc})}{A_{is}(\theta_{is},L^m_{is,soc})}$ v.s. $\frac{B_{is}(\theta_{is},L^m_{is,soc})}{A_{is}(\theta_{is},L^m_{is,soc})}$). The difference in employment between the social optimum and market equilibrium is discussed in section 4.3.2.

Correspondingly, I draw two implications from condition (61): (1) optimal hybrid works' share of time spent working onsite will be larger than market equilibrium if the adjusted relative onsite time elasticity is strong enough (the first effect dominates). (2) When the adjusted relative onsite time elasticity is not strong enough ($\frac{F_{is}(\delta_{is}+1)-\delta_{d,is}}{F_{is}(\delta^R_{is}+1)-\delta^R_{d,is}} < \frac{\zeta_{is}}{e^{kt_i}}$), optimal hybrid works' share of time spent working onsite can also larger than market equilibrium when optimal employment allocation result in relative onsite productivity in socially optimal much higher than that in market equilibrium ($\frac{B_{is}(\theta_{is},L^m_{is,soc})}{A_{is}(\theta_{is},L^m_{is,soc})} > \frac{B_{is}(\theta_{is},L^m_{is,soc})}{A_{is}(\theta_{is},L^m_{is,soc})}$). Condition (62) has similar implications based on the strength of adjusted relative onsite time elasticity and relative onsite productivity.

4.3.2 The Optimal Employment Composition

The socially optimal employment by city-sector-work mode is

$$L^{m*}_{is} = \left(\frac{\widetilde{\phi}^m_{is}}{\widetilde{\widetilde{\Phi}}_i}\right)^{\sigma} \bar{L}_i, \tag{63}$$

where $\widetilde{\phi}^m_{is} = \frac{\widetilde{I}^m_{is}\epsilon^m_{is}}{d^m_{is}(\theta^m_{is})}$, $\widetilde{\widetilde{\Phi}}_i = \left(\sum_s \sum_m \left(\widetilde{\phi}^m_{is}\right)^{\sigma}\right)^{\frac{1}{\sigma}}$. Equation (63) has the same form as the labor supply function in market equilibrium (29). However, the socially optimal expenditure (corresponding to income in the market equilibrium) is different. The socially optimal expenditure for a worker in city i, sector s, work mode m is

$$\widetilde{I}^m_{is} = \frac{\sigma}{\sigma+1}\left(\widetilde{P}\frac{\partial Y_i}{\partial L^m_{is}} - e_i\right), \tag{64}$$

where $\widetilde{P}$ is the shadow price of the tradable goods produced by city i. $\frac{\partial Y_i}{\partial L_{is}^m}$ is the marginal product of workers in city i, sector s, and work mode m. e_i is the opportunity cost of reallocating workers to other sector-work modes within a city. The marginal product of a worker in the social planner problem is:

$$\frac{\partial Y_i}{\partial L_{is}^m} = \frac{\partial Y_i}{\partial y_{is}} \Bigg(y_{is}^m + \underbrace{\sum_{m' \in \{o,h,f\}} \frac{\partial y_{is}^{m'}}{\partial L_{is}^m}}_{\text{productivity spillover}} \Bigg), \tag{65}$$

where $y_{is}^o = B_{is}(L_{is}^m, \theta_{is}), y_{is}^h = \beta_{is}\ell_{is}(L_{is}^m, \theta_{is}), y_{is}^h = A_{is}(L_{is}^m, \theta_{is})$. Equation (F.50) shows that an increase of employment in city i sector s work mode m affects the total product of city i through 3 channels: (1) the direct effect for sector s work mode m (y_{is}^m): more workers in the sector-work mode cell result in increased production for this sector and work mode. (2) the productivity spillover within sector s ($\sum_{m'} \frac{\partial y_{is}^{m'}}{\partial L_{is}^m}$): for example, if the number of onsite workers in sector s increases, onsite productivity also increases. In the presence of productivity spillover between onsite and remote workers ($\tau > 0$), remote productivity also increases. The increase in onsite and remote productivity boosts hybrid workers' productivity. (3) the marginal effect of the sector s on city-wide productivity ($\frac{\partial Y_i}{\partial y_{is}}$): more workers in sector s means and the total productivity of the city more rely on sector s. With a fixed city size, it also means fewer workers and decreased productivity in other sectors. The social planner chooses a labor composition where the benefit of increasing workers in sector s exceeds the productivity loss in other sectors.

The opportunity cost of reallocating workers to other sector-work modes within a city is

$$e_i = \widetilde{P} \sum_s \sum_m \left(\frac{L_{is}^m}{\bar{L}_i} \frac{\partial Y_i}{\partial L_{is}^m} \right) - \frac{\sigma + 1}{\sigma} \sum_s \sum_m \frac{L_{is}^m}{\bar{L}_i} \widetilde{I}_{is}^m, \tag{66}$$

where the first term is the weighted average of the value of the marginal product of labor across all sectors and work modes in a city. The weight is the share of employment in a sector-work mode as a percentage of total employment in a city. The second term is the average wage adjusted by labor supply elasticity.

Optimal expenditure (F.48) differs from market equilibrium income ($I_{is}^m = (1 +$

$\bar{r})P\frac{\partial Y_i}{\partial y_{is}}y^m_{is})$ by incorporating the productivity spillover effect and the opportunity cost of reallocation. These are also two channels through which the socially optimal employment composition differs from the market equilibrium, as the employment size of a sector-work mode is positively correlated with expenditure. In other words, the social planner allocates more workers to a sector-work mode compared to market equilibrium if it generates more productivity spillover and is associated with a low opportunity cost.

5 Data and Structural Estimation

5.1 Data

I use U.S. data to perform structural estimation, quantify the gap between the social optimum and the market outcome, and measure optimal subsidies.

Dataset, Time, and Samples. I derive the average wage, employment, and share of time hybrid workers spend working onsite using data from the IPUMS Current Population Survey (CPS) basic monthly survey from 2022 Oct to 2024 December. The samples consist of labor forces excluding self-employed, armed forces, and unpaid family workers. I calculate the average commuting time using data from the 2022 American Community Survey in the IPUMS USA dataset. The samples for calculating average commuting time consist of workers who live and work in the same city.

Sectors and Cities. I match the sectors to 13 time-consistent sectors classified by Pollard (2019)[12] I match the cities in the model to the Core-Based Statistics Areas (CBSAs). I map the available geographic units of CPS and ACS to the 2015 version of CBSA. For CPS data, I match the resident metropolitan area or county to CBSA using the crosswalk in the Missouri Census Data Center (MCDC)[13]. If one resident city belongs to multiple CBSAs, the samples are duplicated, and each duplicate is assigned an allocation factor that denotes the share of the original location in the matched CBSA.

[12] The sectors are Agriculture, forestry, fishing, and hunting, Mining, Construction, Manufacturing, Wholesale and retail trade, Transportation and utilities, Information, Financial activities, Professional and business services, Educational and health services, Leisure and hospitality, Other services, and Public administration.

[13] Website address for the MCDC crosswalk files is https://mcdc.missouri.edu/applications/geocorr.html.

The allocation factor is multiplied by the sample weight when calculating the average variables.

For ACS data, I first match place of work public use micro areas (PWPUMAs) to public use micro areas (PUMAs) and then match PUMAs to CBSAs. I use the PW-PUMAs to PUMAs crosswalk from IPUMS and the PUMAs to CBSAs crosswalk from MCDC [14]. If one PWPUMA or PUMA belongs to multiple CBSAs, it is matched to the CBSA with the largest proportion of the population of the PWPUMA.[15]

Hybrid Workers' Share of Time Spent Working Onsite. The share of time spent working onsite is calculated by one minus the share of time WFH. CPS basic monthly survey variable TELWRKHR reports hours WFH for pay in the last week. I only include workers who have one job (identified by the variable MULTJOB) to calculate the average share of time spent WFH. Share of time spent WFH is calculated by hours WFH per week divided by hours worked per week. If the workers report not WFH (reported in the variable TELWRKPAY) or their WFH hours are zero, they are classified as onsite workers. If workers' weekly WFH hour equals their working hours, they are classified as fully WFH workers.

Residual Wages. I use average residual wages to match wages in the model and recover the shifters. For city-sector cells with positive employment but missing wages, I impute the wages using fitted values from regressions of average residual wages with fixed effects. Data process details are in Appendix H.

Employment. I use the CPS basic monthly survey to calculate employment. Distinguishing work modes requires information about the share of time spent WFH. The observations that report the share of time spent WFH are a subset of the basic monthly survey. Therefore, I calculate the employment in the following steps. Firstly, I use a whole set of basic monthly survey data to calculate the monthly average employment

[14] One PWPUMAs may contain several PUMAs. The place of work PUMAs to PUMAs crosswalk are at https://usa.ipums.org/usa/volii/00pwpuma.shtml and https://usa.ipums.org/usa/volii/10pwpuma.shtml.

[15] I choose ACS samples who live and work in the same CBSA. Keeping the matches for multiple CBSAs and applying an allocation factor like in CPS data would require extra assumptions about workers' residence-workplace pairs, which I want to avoid. Thus, I match 1 PWPUMA or PUMA to one CBSA.

by city and sector. Then, I use the subset data that reports the share of time spent WFH to calculate onsite, hybrid, and fully WFH employment shares by city-sector cells. Finally, the employment for each city-sector-work mode is calculated by multiplying the employment in each city-sector cell by the respective employment share of the work mode.

5.2 Identification

I apply generalized method of moments (GMM) to jointly estimate the extensive margins of onsite and remote productivity externalities, cross-worksite spillover, and elasticity of substitution between WFH and working onsite (section 5.3). The method is inspired by Ahlfeldt et al. (2015) and Farrokhi (2021), leveraing orthogonal conditions for identification. Although the parameters are estimated jointly, each moment condition is more closely associated with identifying a specific parameter. This section discusses the identification intuition behind how each moment condition helps identify the corresponding parameter.

5.2.1 Extensive Margins of Remote and Onsite Productivity Externalities

The variation of fully WFH wages in response to the fully WFH employment identifies the remote agglomeration elasticity λ^R. According to the model, fully WFH wages can also be expressed as $W_{is}^f = w_{is}(\bar{y}_{is})\bar{a}_{is} f(\frac{L_{is}^h}{L_{is}^f}, \frac{L_{is}^o}{L_{is}^f}, \theta_{is})(\bar{L}_{is}^f)^{\lambda^R}$, where $w_{is} = P\bar{y}_{is}\left(\frac{Y_i}{y_{is}}\right)^{\frac{1}{\eta}}$, $f(\frac{L_{is}^h}{L_{is}^f}, \frac{L_{is}^o}{L_{is}^f}, \theta_{is}) = \left[1 + (1-\theta_{is})\frac{L_{is}^h}{L_{is}^f} + \tau(\frac{L_{is}^o}{L_{is}^f} + \theta_{is}\frac{L_{is}^h}{L_{is}^f})\right]^{\lambda^R}$.[16] The log form of this equation implies the following labor demand regression:

$$ln(W_{is}^f) = \lambda^R ln(L_{is}^f) + x_{is}^f + ln(\bar{y}_{is}\bar{a}_{is}), \tag{67}$$

where $x_{is}^f = ln(P(\frac{Y_i}{y_{is}})^{\frac{1}{\eta}}) + ln(f(\frac{L_{is}^o}{L_{is}^f}, \frac{L_{is}^h}{L_{is}^f}, \theta_{is}))$. Equation (67) implies that a stronger remote productivity externality (λ^R) leads to an increase in fully WFH wages, given other things

[16] This equation describes the demand for fully remote workers. It captures two types of relationships between fully remote labor and wages: a negative relationship at the city-sector level, represented by the labor demand elasticity η, and a positive relationship at the city-sector–work mode level, represented by the remote agglomeration elasticity λ^R. When there is no remote productivity spillover ($\lambda^R = 0$), the wage for fully remote workers equals the city-sector wage (w_{is}) and is negatively related to fully remote employment. In contrast, if there is a positive remote productivity spillover ($\lambda^R > 0$), then, holding the city-sector wage constant, the fully remote workers' wage increases with the level of fully remote employment within that city-sector.

unchanged. The variation of fully WFH wages in response to fully WFH employment across city-sector cells identifies the remote agglomeration elasticity.

Since the exogenous city-sector-specific productivity ($\bar{y}_{is}$) and the exogenous remote productivity ($\bar{a}_{is}$) are the residuals of the regression (67), the identification moment condition in a simple OLS regression is $E(L^f_{is}, \bar{y}_{is}\bar{a}_{is}) = 0$. However, the OLS regression has three identification issues: (1) reverse causality: higher remote wages in some city-sector cells attract more workers to choose these remote jobs, rather than more remote workers leading to higher fully WFH wages. (2) endogeneity: the residuals may correlate with the remote employment. For example, larger cities tend to have higher wages and more remote workers. This implies a positive correlation between residual and remote employment, which leads to an upward bias of remote productivity externality estimates. (3) selection bias: people with different productivity sorting into remote jobs. Harrington and Emanuel (2021) find that call center workers with less productivity are more likely to choose remotely than onsite. This negative selection implies a downward bias of the estimate of remote productivity externality. On the other hand, if highly productive workers choose fully WFH, it will lead to a positive bias.

To obtain an unbiased estimate of λ^R, we may use an instrumental variable that is correlated with fully remote employment but uncorrelated with the residuals of fully remote demand ($\bar{y}_{is}\bar{a}_{is}$). For identification, I impose the orthogonality condition that the fully remote labor supply shifter (ϵ^f_{is}) is uncorrelated with the residuals of fully remote demand:

$$E(\epsilon^f_{is}, \bar{y}_{is}\bar{a}_{is}) = 0 \tag{68}$$

Given the labor supply elasticity, the GMM method estimates the remote productivity externality by minimizing the correlation between model-implied supply and demand shifters for fully remote labor. When the method yields an estimate that is consistent with the moment condition (i.e., the correlation between these shifters is nearly zero), the approach is similar to treating the fully remote labor supply shifters as an instrumental variable for fully remote employment.

Hybrid workers' wages can also used to identify remote productivity externality. In the model, the ratio of hybrid wages to onsite wages is:

$$\frac{W_{is}^h}{W_{is}^o} = \beta_{is}\left[\left(\frac{\bar{a}_{is}}{\bar{b}_{is}}L_{is}^{\lambda^R-\lambda}x_{is}^h\right)^{\frac{\rho-1}{\rho}} + \theta_{is}^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}, \tag{69}$$

where $x_{is}^h = g(\pi_{is}^o, \pi_{is}^h, \pi_{is}^f, 1-\theta_{is})$ is a function of labor composition and hybrid workers' share of time spent WFH. π_{is}^m ($m \in \{o, h, f\}$) denotes the percentage of workers in work mode m relative to the total employment in the city-sector. Equation (69) suggests that given onsite productivity externality (λ), a stronger remote productivity externality (λ^R) leads to a higher increase in hybrid wage ratio relative to total employment, holding other things unchanged. Given the value of onsite productivity externality, the variation of the hybrid wage ratio across city-sector cells in response to employment identifies the remote productivity externality. To deal with endogenous issues, I leverage the orthogonal assumption that the labor supply shifter is uncorrelated with the residuals of relative hybrid labor demand:

$$E(\epsilon_{is}, \frac{\beta_{is}\bar{a}_{is}}{\bar{b}_{is}}) = 0, \tag{70}$$

where $\frac{\beta_{is}\bar{a}_{is}}{\bar{b}_{is}}$ is the residuals in equation (69), which contains the exogenous remote productivity relative to onsite productivity ($\frac{\bar{a}_{is}}{\bar{b}_{is}}$) and exogenous hybrid productivity (β_{is}). $\epsilon_{is} = (\sum_m \epsilon_{is}^m)^\sigma$ combines labor supply shifter for three work-modes in a city-sector cell (equation (31) shows how this shifter is constructed based on labor supply from three work-modes). When the GMM yields an estimate of λ^R that is consistent with the moment condition, the approach is akin to treating the city-sector level labor supply shifters as an instrumental variable for employment at the city-sector level.

Similar to (67), the variation of onsite wages in response to the onsite employment identifies the onsite agglomeration elasticity λ. The onsite wages in the model is $W_{is}^o = w_{is}(\bar{y}_{is})\bar{b}_{is}f\left(\frac{L_{is}^h}{L_{is}^o}, \frac{L_{is}^f}{L_{is}^o}, \theta_{is}\right)(L_{is}^o)^\lambda$, which implies the following regression:

$$ln(W_{is}^o) = \lambda ln(L_{is}^o) + x_{is}^o + ln(\bar{y}_{is}\bar{b}_{is}), \tag{71}$$

where $x_{is}^o = ln(P(\frac{Y_i}{y_{is}})^{\frac{1}{\eta}}) + ln(f(\frac{L_{is}^h}{L_{is}^o}, \frac{L_{is}^f}{L_{is}^o}, \theta_{is}))$. Equation (71) implies that a stronger onsite productivity externality (λ) leads to an increase in onsite wages, given other things unchanged. The variation of onsite wages in response to onsite employment across city-sector cells identifies the onsite agglomeration elasticity. The moment condition for identifying onsite agglomeration elasticity is:

$$E(\epsilon^o_{is}, \bar{y}_{is}\bar{b}_{is}) = 0, \tag{72}$$

where $\bar{y}_{is}\bar{b}_{is}$ is the residual in equation (71). Given labor supply elasticity, when the estimate of λ makes the onsite labor supply shifter uncorrelated with the residuals of onsite demand, the approach is similar to using the onsite labor supply shifter as an instrumental variable for onsite employment.

5.2.2 Cross-worksite Spillover

This section explains the intuition of estimating cross-worksite spillover. The variation of relative onsite productivity across city-sector cells in response to the share of time spent working onsite identifies the cross-worksite spillover.[17]

For simplicity, consider the relative demand curve for the onsite time in the basic model:

$$\frac{B}{A} = f(\theta, \tau) = \frac{\bar{b}}{\bar{a}} \frac{(\theta + \tau(1-\theta))^{\lambda}}{(1-\theta+\tau\theta)^{\lambda^R}} L^{\lambda-\lambda^R} = \underbrace{\frac{\bar{b}}{\bar{a}}}_{\text{shifter}} g(\theta, \tau) L^{\lambda-\lambda^R}. \tag{73}$$

The cross-worksite spillover τ affects the slope of this relative demand curve. When $\tau = 0$, relative onsite productivity increases in the share of time spent working onsite. When $\tau = 1$, relative onsite productivity is irrelevant to the share of time spent working onsite, implying a horizontal relative demand curve. As τ increases from 0 to 1, the relative onsite productivity increases less given the same level increase in the share of time spent working onsite, implying a decrease in the slope of the relative demand curve.

In the full model, relative onsite productivity and share of time spent working onsite vary by city-sector cells. Therefore, if there are no endogenous issues, the city-sector

[17] A simple way to obtain relative onsite productivity is using the ratio of onsite wages to fully WFH wages ($\frac{B_{is}}{A_{is}} = \frac{W^o_{is}}{W^f_{is}}$). Appendix I.1 also describes the procedure for recovering remote productivity using hybrid wages.

variation of the relative onsite productivity in response to differences in the share of time spent working onsite identifies the cross-worksite spillover. It implies an identification moment condition $E(\theta_{is}, \frac{\bar{b}_{is}}{\bar{a}_{is}}) = 0$, where exogenous relative onsite productivity is the residual/shifter of the relative demand curve.

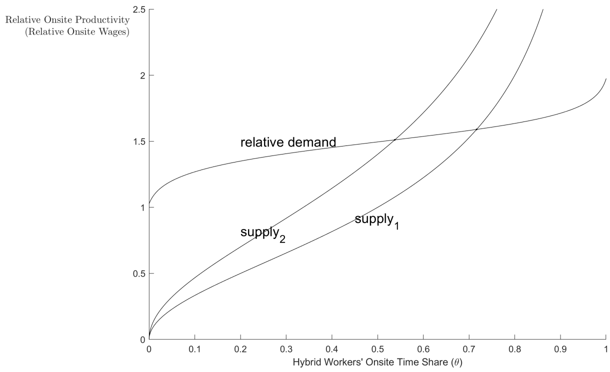


Figure 10: Share of Time Working Onsite and Relative Onsite Productivity

However, the share of time spent working onsite and the relative onsite productivity are simultaneously determined. In the market equilibrium, workers' choice of the share of time spent working onsite is given by:

$$\theta = f(\frac{B}{A}) = \frac{1}{1 + \left(\frac{B}{A}\right)^{1-\rho} \underbrace{\left(\frac{\zeta}{\zeta^o}\right)^{-\rho}}_{\text{shifter}}} \tag{74}$$

This supply curve implies higher relative onsite productivity leads to a higher share of time spent working onsite. The relative onsite amenity $\frac{\zeta}{\zeta^o}$ is a shifter for the supply of the onsite time. The simultaneous equations for the onsite time and the relative onsite productivity imply the condition $E(\theta_{is}, \frac{\bar{b}_{is}}{\bar{a}_{is}}) = 0$ suffers from an endogenous issue. To address this issue, I use the relative amenity shifters from the supply side to identify the cross-worksite spillover. Figure 10 illustrates this concept that a supply side shifter can identify the slope of the relative demand curve. The moment condition for estimating τ is:

$$E(\frac{\zeta_{is}}{e^{kt_i}}, \frac{\bar{b}_{is}}{\bar{a}_{is}}) = 0, \tag{75}$$

When the estimate of τ yields model-recovered shifters consistent with the moment condition, the approach is similar to using the relative amenity shifter as an instrumental variable for the share of time spent working onsite.

5.2.3 Elasticity of Substitution between Woring Onsite and WFH

The first-order condition with respect to hybrid workers' relative share of time spent WFH is $\frac{1-\theta_{is}}{\theta_{is}} = \left(\frac{A_{is}}{B_{is}}\right)^{\rho-1} (\frac{e^{kt_i}}{\zeta_{is}})^{\rho}$. Taking the log of both sides of the equation yields

$$ln(\frac{1-\theta_{is}}{\theta_{is}}) = \rho k t_i + x_{is} + \epsilon_{is}(\frac{\bar{a}^s}{\bar{b}_{is}}, \zeta_{is}), \tag{76}$$

where $x_{is} = (\rho - 1)ln\left(\frac{(L^A_{is}(\theta_{is})+\tau L^B_{is}(\theta_{is}))^{\lambda T}}{\left(L^B_{is}(\theta_{is})+\tau L^A_{is}(\theta_{is})\right)^{\lambda}}\right)$, $\epsilon_{is}(\frac{\bar{a}^s}{\bar{b}_{is}}, \zeta_{is}) = (\rho-1)ln\frac{\bar{a}^s}{\bar{b}_{is}} + \rho ln(\zeta_{is})$ is the residual containing the exogenous relative WFH productivity $(\frac{\bar{a}^s}{\bar{b}_{is}})$ and WFH amenity costs (ζ_{is}). Given the elasticity of commuting cost k, a higher elasticity of substitution between working onsite and WFH (ρ) implies that hybrid workers' relative WFH time increases more as commuting time becomes longer. The variation of hybrid workers' relative share of time spent WFH in response to the commuting time identifies ρ. The moment condition for identifying the elasticity of substitution between working onsite and WFH is:

$$E(t_i, \zeta_{is}\frac{\bar{a}^s}{\bar{b}_{is}}) = 0. \tag{77}$$

5.3 Joint Estimation

λ, λ^R, τ, and ρ are jointly estimated using GMM approach. They are the solution to the problem:

$$\boldsymbol{\beta} = argmin_{\boldsymbol{\beta}} \boldsymbol{g(\beta)'Dg(\beta)}, \tag{78}$$

$$\text{s.t. } \bar{\boldsymbol{X}} = \mathscr{F}(\boldsymbol{X_{data}}; \boldsymbol{\beta}, \bar{\boldsymbol{\beta}}). \tag{79}$$

$$\boldsymbol{X} = \mathscr{G}(\bar{\boldsymbol{X}}; \boldsymbol{\beta}, \bar{\boldsymbol{\beta}}). \tag{80}$$

The vector $\boldsymbol{\beta} = [\lambda, \lambda^R, \tau, \rho]'$, $\boldsymbol{D}$ is a diagonal matrix. $\boldsymbol{g(\beta)}$ consists of differences between moments predicted by the model and the target moments.

$$\boldsymbol{g(\beta)} = \begin{bmatrix} \text{weighted } cov(\epsilon^o_{is}, \bar{y}_{is}\bar{b}_{is}) - 0 \\ \text{weighted } cov(\epsilon^f_{is}, \bar{y}_{is}\bar{a}_{is}) - 0 \\ \text{weighted } cov(\epsilon_{is}, \beta_{is}\frac{\bar{a}_{is}}{\bar{b}_{is}}) - 0 \\ \text{weighted } cov(\frac{\zeta_{is}}{e^{kt_i}}, \frac{\bar{b}_{is}}{\bar{a}_{is}}) - 0 \\ \text{weighted } cov(t_i, \zeta_{is}\frac{\bar{a}_{is}}{\bar{b}_{is}}) - 0 \end{bmatrix}, \tag{81}$$

where weighted $cov(x_{is}, y_{is}) = \sum_i \sum_s (wgt_{is} \cdot (x_{is} - \bar{x}_{is}) \cdot (y_{is} - \bar{y}_{is}))$. The weight is the share of city-sector employment as a percentage of total employment: $wgt_{is} = \frac{L_{is}}{\sum_i \sum_s L_{is}}$, $\bar{x}_{is} = \sum_i \sum_s (wgt_{is} \cdot x_{is})$, $\bar{y}_{is} = \sum_i \sum_s (wgt_{is} \cdot y_{is})$

$\mathscr{F}(\boldsymbol{X}; \boldsymbol{\beta}, \bar{\boldsymbol{\beta}})$ is the model inversion function described in Appendix I.1. Given parameter vector $\boldsymbol{\beta}$ and $\bar{\boldsymbol{\beta}} = [\sigma, \eta, k]'$, this function uses data $\boldsymbol{X}$ to recover model shifters $\bar{\boldsymbol{X}}$. Variables from data include wages, employment, the share of time spent working onsite, and commuting time ($\boldsymbol{X} = [L^m_{is}, W^m_{is}, \theta_{is}, t_i], m \in \{o, h, f\}$). The vector of shifters includes labor demand and supply shifters and WFH amenity cost ($\bar{\boldsymbol{X}} = [\bar{y}_{is}, \bar{b}_{is}, \bar{a}_{is}, \beta_{is}, \epsilon^o_{is}, \epsilon^h_{is}, \epsilon^f_{is}, \zeta_{is}]$). The values of the parameters in $\bar{\boldsymbol{\beta}}$ are borrowed from literature. $\mathscr{G}\left(\bar{\boldsymbol{X}}; \boldsymbol{\beta}, \bar{\boldsymbol{\beta}}\right)$ solve for endogenous variables in the model given the shifters and parameters (Appendix I.2). Table 10 and 7 in the appendix show the summary statistics of variables from data and recovered shifters. Table 2 summarizes estimate results and other parameters that are given in structural estimation. Table 8 in the appendix shows the error for moment conditions.

Table 2: GMM Parameters

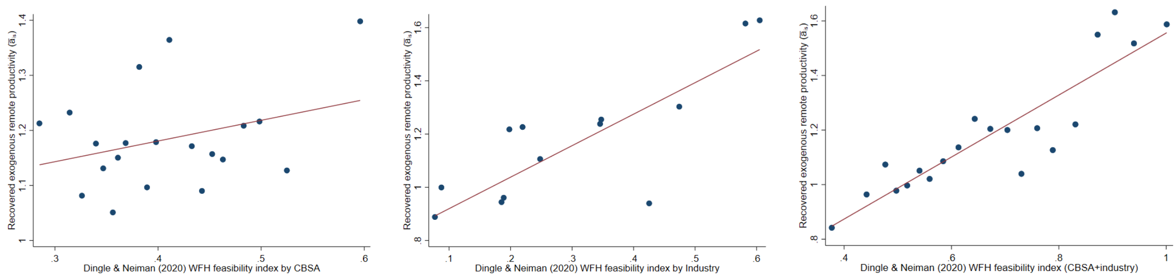
	Description	Value	Moment conditions/Literature
λ	Onsite agglomeration elasticity	0.088	$cov(\varepsilon^o_{is}, \bar{y}_{is}\bar{b}_{is}) = 0$
λ^R	Remote agglomeration elasticity	0.069	$cov(\varepsilon^f_{is}, \bar{y}_{is}\bar{a}_{is}) = 0$ and $cov(\varepsilon_{is}, \beta_{is}\frac{\bar{a}_{is}}{\bar{b}_{is}}) = 0$
τ	Cross-worksite productivity spillover	0.009	$cov(\frac{\zeta_{is}}{e^{kt_i}}, \frac{\bar{b}_{is}}{\bar{a}_{is}}) = 0$
ρ	Elasticity of substitution between WFH and onsite work	1.304	$cov(t_i, \frac{\zeta_{is}\bar{a}_{is}}{\bar{b}_{is}}) = 0$
σ	Labor supply elasticity	1.26	Burstein et al. (2019)
η	Labor demand elasticity	0.8	Lichter et al. (2015),Beaudry et al. (2018)
k	Elasticity of commuting costs	0.01	Ahlfeldt et al. (2015)

Note: All covariances in this table are weighted by the city-sector cells' employment.

The result shows that the estimated onsite agglomeration externality is larger than the remote agglomeration externality. The difference between onsite and remote elasticities is around 0.02, which is within the range of the estimates obtained from the IV regressions in Appendix K. The onsite and remote elasticities are also comparable to the physical agglomeration elasticity estimated in the literature. The urban agglomeration elasticity ranges from -0.09 to 0.19 in Melo et al. (2009), 0.04 to 0.07 in Combes and Gobillon (2015), and 0.02 to 0.05 in Rossi-Hansberg et al. (2021). The estimated workplace agglomeration elasticity is 0.07 in Ahlfeldt et al. (2015) and 0.086 in Heblich et al. (2020).

The estimated cross-worksite spillover is 0.009. It suggests a substantial efficiency loss when sharing knowledge between onsite and remote workers. The estimated elasticity of substitution between WFH and onsite work is around 1.3. It is within 95 % confidence interval range of the elasticity of substitution between working onsite and WFH estimated by Davis et al. (2024) (0.998 to 6.105). It is smaller than the education-sector-specific elasticity of substitution between WFH and onsite work calibrated by Delventhal and Parkhomenko (2023) (3.0 to 4.3).

5.4 Model Fit



(a) (b) (c)

Figure 11: Exogenous WFH Productivity and WFH Feasibility Index

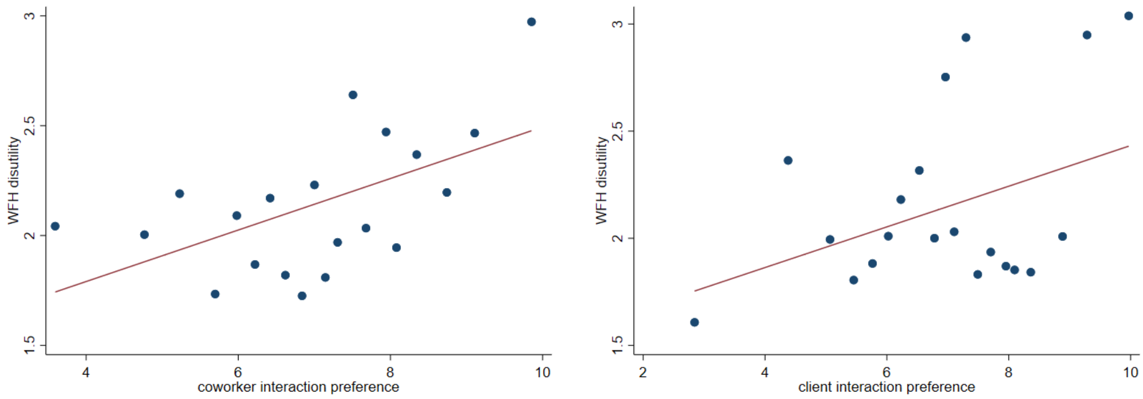
Note: The X-axis in the three subfigures shows the WFH feasiblity index based on Dingel and Neiman (2020). The Y-axis represents the recovered exogenous remote productivity by CBSA and industry.

Figure 11 shows the positive correlation between recovered exogenous remote productivity at the CBSA-industry level and the WFH feasiblity index constructed by Dingel and Neiman (2020). The X-axis in Figure 11(a) shows Dingel and Neiman (2020)'s index at the metropolitan statistical area (MSA) level[18]. In Figure 11(b), I match Dingel

[18] CBSAs consist of metropolitan statistical areas (MSA) and micropolitan statistical areas (μSA).

and Neiman (2020)'s industry index to the industry classification in this paper (Pollard (2019)). In Figure 11(c), the X-axis is the sum of the MSA level and industry level indexes in the previous two figures. Figure 11 indicates that the recovered remote productivity shifter captures the variation in WFH feasibility across cities and industries.

According to the Survey of Working Arrangements and Attitudes (SWAA), the top two benefits of working onsite are "face-to-face collaboration" and "socializing". This indicates that one of the main disutilities of WFH is the lack of in-person interactions. I obtain measures of in-person interaction preference from the SWAA at the combined statistical area (CSA)-industry level and then compare them with the recovered WFH disutility in Figure 12. The x-axes of the two subfigures correspond to preferences for in-person interaction with coworkers and clients, respectively. A higher number indicates a greater enjoyment of in-person interaction. The binscatter plots in Figure 12 show that region-industry cells with a higher preference for in-person interactions with coworkers and clients tend to have a larger WFH disutility. This suggests that the recovered WFH disutility correlates with the lack of in-person interaction.



(a) (b)

Figure 12: WFH Disutility and In-person Interaction Preference

Note: The X-axes in the two subfigures represent workers' preferences for in-person interactions with coworkers or clients. The data source is responses to two questions from the SWAA (August–October 2021): "How much do you enjoy personal interactions with coworkers at your employer's worksite?" and "How much do you enjoy personal interactions with customers, clients, or patients at your employer's worksite?". Respondents rated their enjoyment on a scale from 1 ("not at all") to 10 ("very much"). The data are aggregated by combined statistical area (CSA) and industry. The Y-axis shows the recovered work-from-home (WFH) disutility at the CSA-industry level.

6 Quantification Results

Given the estimated parameters and shifters, I solve for market equilibrium and the socially optimal allocation to quantify the gap between them, as well as the socially optimal subsidies. The algorithm for solving market equilibrium is described in Appendix I.2. The solution for the social planner's problem is in Appendix E. The socially optimal policies are explained in Appendix G. Appendix J shows the parameters and shifters used for the quantification analysis. The following sections present the gap and subsidies at the average level and then analysis the variations of intensive and extensive margins of labor at the city-sector level.

6.1 The Gap between the Social Optimum and the Market Equilibrium, and Subsidies at the Aggregate Level

Table 3: Gaps and Subsidies

	Exist remote productity spillover	No remote productity spillover ($\lambda^R = \tau = 0$)
Welfare gain in social optimum, % chg	2.27	2.25
Share of time hybrid workers spend working onsite, % chg	2.87	5.43
Onsite employment, % chg	1.75	2.55
Hybrid employment, % chg	-4.91	-5.22
Fully WFH employment, % chg	-4.87	-7.42
Share of onsite subsidy in hybrid workers' gross income (income tax), %	10.98	15.43
Share of onsite subsidy in hybrid workers' gross income (WFH tax), %	1.61	2.80

Note: The welfare gain is calculated as $(\frac{U^{soc}}{U^{mkt}} - 1) \times 100\%$, where U^{soc} and U^{mkt} denote social welfare in the social optimum and market equilibrium, respectively. Social welfare is defined by equation (53), which aggregates the expected utility of all cities $U = \sum_i U_i \bar{L}_i$. Other variables represent the weighted average of the gap between the social optimum and the market outcome, with the weights based on the share of city-sector employment in market equilibrium as a percentage of total employment. For example, let θ_{is}^{soc} and θ_{is}^{mkt} refer to the share of time spent working onsite in city i, sector s in the social optimum and market equilibrium, respectively. The value in the table is calculated as $\frac{\sum_i \sum_s \Delta\theta_{is} \times L_{is}^{mkt}}{\sum_i \sum_s L_{is}^{mkt}}$, where $\Delta\theta_{is} = \left(\frac{\theta_{is}^{soc}}{\theta_{is}^{mkt}} - 1\right) \times 100\%$ and L_{is}^{mkt} is employment in city i, sector s in the market equilibrium. Share of onsite subsidy is calculated according to Appendix G.

Table 3 shows the gap between the social optimum and the market outcome and the shares of onsite subsidies with two tax sources. Column 1 shows the results in the presence of remote productivity spillover. Social welfare is around 2.3 % higher than the market equilibrium. The gap for hybrid workers' shares of time spent working onsite and employment are the weighted average across city-sector cells. For the intensive margin of onsite work, the average optimal share of time hybrid workers spend working onsite is around 3 % higher than the market equilibrium. In terms of the extensive margin of labor, average optimal onsite employment is around 2 % higher than market equilibrium, while average hybrid or fully WFH employment is around 5 % lower than market equilibrium.

The policies target optimal hybrid workers' shares of time spent working onsite to achieve the social optimum. By adjusting the intensive margin of labor, the economy also achieves the socially optimal extensive margin of labor allocation through productivity spillover effects (See Appendix G for derivation details.). If the social planner implements an income tax to fund the subsidies for onsite work, the average cost is around 11 % of hybrid workers' before-tax income. If the social planner levies WFH tax to subsidize onsite work, the average cost is around 2 % of hybrid workers' before-tax income. Using the externality tax to fund the subsidies costs less than using the income tax.

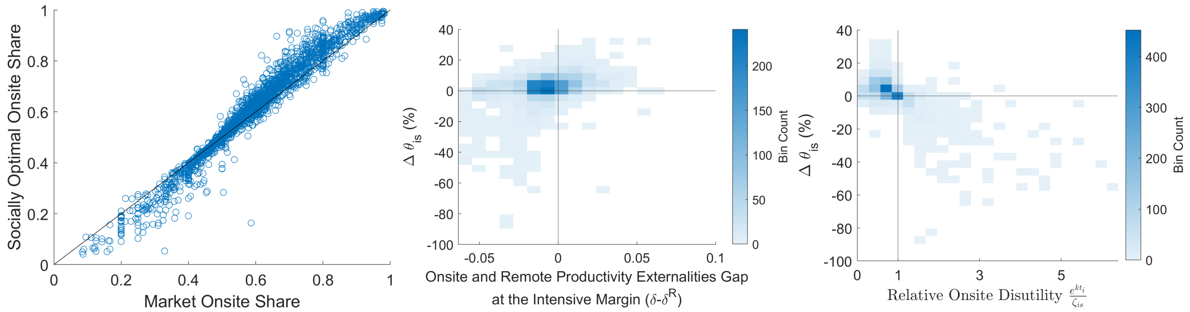
For column 2 in Table 3, I assume there is no remote productivity spillovers by setting remote agglomeration elasticity and the cross-worksite spillover to be zero ($\lambda^R = \tau = 0$). The values of other parameters remain the same as the estimation results. I then recovered the shifters and solve for market equilibrium and socially optimal allocation under this setup. In this scenario, workers compare exogenous remote productivity relative to onsite productivity with spillover effect and relative WFH disutility to decide whether and how much working onsite. Compared to the gap in the presence of remote productivity spillover, column 2 shows that the gaps between the social optimum and the market outcome enlarge in the absence of remote productivity spillover. The socially optimal levels of hybrid workers' onsite time and onsite employment are around 5% and 3% higher than the market equilibrium, respectively. The subsidy required to achieve social optimal become costly: subsidy funding by income tax equals to 15 % of hybrid workers' gross income, subsidy funding by externality tax equals to around 3 % of hybrid workers' gross income.

Table 3 implies that, on average, the social optimum favors more onsite work. However, the remote productivity spillover allows us to get closer to the optimal level of onsite work by compensating for some of the reduced onsite productivity spillover.

6.2 The Variations of the Gaps for Intensive and Extensive Margins of Labor

Although Table 3 shows that optimal allocation features more onsite work at intensive and extensive margins on average, there are variations at city-sector levels. This section discusses the variations of the social optimum-market gap for hybrid workers' shares of time spent working onsite and labor composition for three work modes. I begin with the patterns of the socially optimal allocation and the gap, and then explain the mechanisms driving these patterns.

6.2.1 Hybrid Workers' Share of Time Working Onsite



(a) (b) (c)

Figure 13: Pattern and Determinants of the Gap in Hybrid Workers' Share of Time Working Onsite

Note: The gap between the social optimum and market equilibrium hybrid workers' shares of time spent working onsite is calculated as $\Delta\theta_{is} = (\frac{\theta_{is}^{soc}}{\theta_{is}^{mkt}} - 1) \times 100\%$.

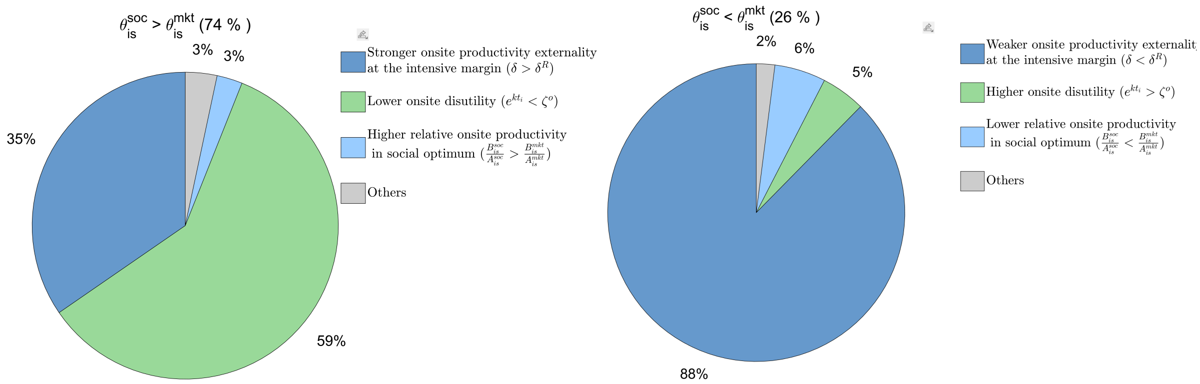
Figure 13(a) shows the S-shaped relationship between optimal and market equilibrium hybrid workers' shares of time spent working onsite. Each point represents a city-sector cell. The figure shows that for city-sector cells where hybrid workers' shares of time spent working onsite in market equilibrium are larger than 50 %, optimal shares of time spent working onsite tend to be higher than market equilibrium. On the contrary, for city-sector cells where hybrid workers' shares of time spent working onsite in market equilibrium are smaller than 50 %, optimal shares of time spent working onsite tend to be smaller than market equilibrium.

As analyzed in section 4.3.1, whether optimal share of time hybrid workers spend working onsite is larger than the market equilibrium depends on intensive margins of productivity externalities (δ and δ^R), other sources of externalities, and relative onsite productivity in the social optimum compared to market equilibrium.

Figures 13(b) show how the gap between onsite and remote productivity externalities at the intensive margins affects the gap between optimal and market equilibrium shares of time spent working onsite. When the intensive onsite productivity externality is stronger than the intensive remote productivity externality, the socially optimal share of time spent working onsite is more likely to be larger than that in market equilibrium. Conversely, when the intensive onsite productivity externality is weaker than the intensive remote productivity externality, the socially optimal share of time spent working onsite is more likely to be smaller than that in market equilibrium. The gap between optimal and market equilibrium shares of time spent working onsite tends to be larger as the gap between intensive onsite and remote productivity externalities becomes larger.

Figure 13(c) shows an overall negative relationship between the gap of optimal and market equilibrium hybrid workers' shares of time spent working onsite and relative onsite disutility (relative WFH amenity). City-sector cells where optimal share of time hybrid workers spend working onsite are smaller than market equilibrium tend to have an onsite disutility larger than WFH amenity costs. By intuition, when onsite disutility is larger than WFH amenity cost, a decrease in onsite work results in a lower composite of disutility for onsite and WFH. Therefore, less onsite time is associated with a lower composite of disutility for hybrid workers in those city-sector cells. This pattern is consistent with the condition (62). A higher relative onsite disutility implies that the intensive onsite productivity externality (δ_{is}) is weakened by a larger discount from the composite of disutility ($\delta_{d,is}$), resulting in a lower adjusted onsite productivity externality relative to adjusted remote productivity externality and a lower optimal onsite share.

To further decompose the factors influencing whether the socially optimal share of time spent working onsite is larger than the market equilibrium, pie charts 14(a) and 14(b) show the share of city-sector cells affected by different forces in two cases. The text above the pie charts shows that 74 % of city-sector cells have an optimal share of time spent working onsite larger than the market equilibrium, while 26 % is the opposite.[19] This pattern is reflected in the average higher optimal share of time spent working onsite in Table 3.



(a) (b)

Figure 14: Decomposition of Factors Influencing the Sign of the Gap

Note: θ_{is}^{soc} denotes the socially optimal hybrid workers' shares of time spent working onsite. θ_{is}^{mkt} denotes the hybrid workers' shares of time spent working onsite in market equilibrium.

Figure 14(a) shows different scenarios for city-sector cells with a higher optimal share of time spent working onsite. Among these cells, 35 % of them have the intensive onsite productivity externality stronger than the intensive remote productivity externality. For 59 % of them, despite having onsite productivity externality weaker than the remote productivity externality at the intensive margin, a lower onsite disutility drives a higher optimal share of time spent working onsite.[20] 3 % of them have a weaker intensive onsite productivity externality and a higher onsite disutility. However, optimal employment results in a higher relative onsite productivity in the social optimum than in market equilibrium, which drives a higher optimal share of time spent working onsite. Figure 14(b) shows that for city-sector cells with a lower optimal share of time spent working onsite, a weaker intensive onsite productivity externality is the main cause.

[19] Although Figure 13 shows that some optimal shares of time spent working onsite are very close to market equilibrium, they are not numerically identical.

[20] When onsite disutility is smaller than WFH amenity costs, an increase in the share of time spent working onsite decreases the disutility composite, leading the social planner to favor a higher share of time spent working onsite.

6.2.2 Employment by Work Mode

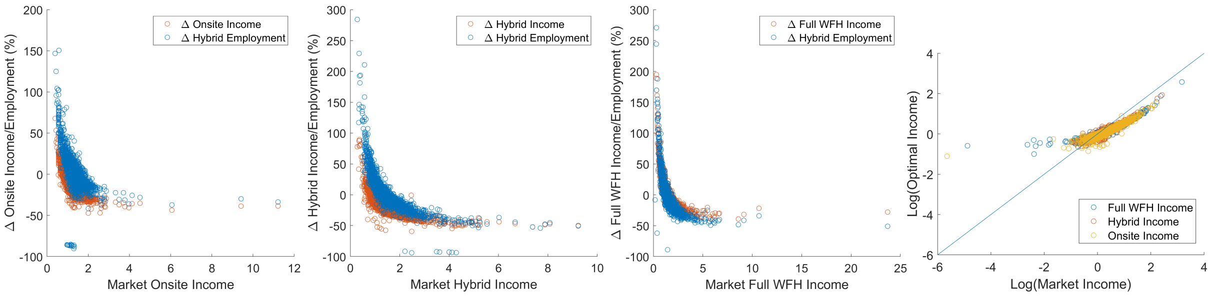


Figure 15: Employment Reallocation by City-Sector-Work mode

Note: Each point represents a city-sector cell. The gap between the social optimum and market equilibrium employment is calculated as $\Delta x_{is} = (\frac{x_{is}^{soc}}{x_{is}^{mkt}} - 1) \times 100\%$.

Although Table 3 shows an increase in onsite employment and decreases in hybrid and full-WFH employment at the aggregate level, the socially optimal allocation has decreased onsite workers and increased hybrid and full-WFH employment at some city-sector cells.

The first three sub-figures of Figure 15 show the reallocation pattern of employment from the social planner's perspective. Compared to the original market equilibrium, the socially optimal allocation reallocates workers from original relatively high-income sectors to relatively low-income sectors. This pattern holds for onsite, hybrid, and fully WFH work modes. As a result, the productivity in original low income sectors increase due to the productivity externalities, leading a increase in income for those sectors. The last sub figure of Figure 15 shows that sectors that original have high income levels tend to retain relatively higher incomes after the reallocation, but the income premium declines.[21]

The employment reallocation pattern is consistent with equation (F.48). According to the estimation results from sector 6, the extensive margins of onsite and remote productivity externalities are smaller than 1, suggesting that the productivity externalities experience diminishing returns. Sectors with original high income tend to have high productivity. Due to diminishing returns, the marginal productivity increase from having more workers in the high-income sector is lower than in the relatively low-income sector.

[21] This pattern is similar to Fajgelbaum and Gaubert (2020). Equation (F.48) in this paper is similar to Equation (88) in Appendix D of Fajgelbaum and Gaubert (2020). Fajgelbaum and Gaubert (2020) considers productivity spillover effect and mobility across cities. The socially optimal employment allocation in their paper features a reallocation from large to small cities and a decrease in urban skill premium.

The productivity spillover effects are stronger in the relatively low-income sectors, which drives the reallocation towards these sectors.

7 Conclusion

This study incorporate productivity spillover effects both within and between onsite and remote labor in a quantitative spatial model. In the model, each worker selects a sector and work mode in each city. Hybrid workers choose the share of time spent working onsite and WFH. I estimate the extensive margins of onsite and remote productivity externalities and cross-worksite spillover using model-implied instrumental variables. The estimation results indicate that the onsite productivity externality is stronger than the remote productivity externality at the extensive margin, and there is a notable efficiency loss in cross-worksite spillover.

The socially optimal allocation differs from the market equilibrium in terms of share of time hybrid workers spend working onsite and employment composition across cities and sectors. The socially optimal share of time hybrid workers spend working onsite depends on the relative onsite time elasticity adjusted by other sources of externalities. A stronger onsite productivity externality relative to the remote productivity externality at the extensive margin does not necessarily guarantee that the socially optimal share of time hybrid workers spend working onsite is higher than the market equilibrium.

I apply the model to U.S. data at the CBSAs and industry levels. The quantification results show that a substantial welfare gain in the socially optimal allocation is associated with a higher share of time hybrid workers spend working onsite and more onsite workers compared to the market equilibrium at the average level. However, there are variations at the city-sector level. City-sector cells with a stronger (weaker) onsite productivity externality relative to remote productivity externality at the intensive margin tend to have a higher (lower) optimal share of time spent working onsite relative to the market equilibrium. In addition, city-sector cells with larger onsite disutility relative to the WFH amenity costs are more likely to have a lower optimal share of time spent working onsite because the externality from the disutility composite weakens the premium of onsite productivity externality at the intensive margin. The gap between optimal and market equilibrium shares of time spent working onsite tends to widen with the gap between the onsite and remote productivity externalities at the intensive margin.

I consider optimal policies targeting the share of time hybrid workers spend working onsite to simultaneously achieve optimal intensive and extensive margins of labor. The quantification results show that implementing an externality tax to fund the subsidies for onsite work is less costly than using income tax.

Several aspects are worth further exploration. First, this paper focuses on hybrid workers and adopts simplified assumptions regarding the colocation of residence and workplace and the stability of city size. However, WFH provides the feasibility of separating residence and workplace, which may lead to migration across cities in the long term (Delventhal and Parkhomenko (2023)). Second, the model assumes the remote productivity externality occurs within each city. However, the remote productivity externality has the potential to extend beyond geographic boundaries. Third, this paper considers socially optimal policies that target the hybrid workers' work arrange. Other optimal policies such as offering transfers across workers to adjust employment allocation could also achieve the socially optimal allocation. Additionally, the model in this paper assumes firms do not make policies about remote work. In the market equilibrium, the number of employments across work modes and hybrid workers' work arrangements is jointly determined by workers' choices and societal productivity levels. However, firms are implementing return-to-office policies as documented by Ding and Ma (2023) and Flynn et al. (2024). These policies may only partially and locally internalize the externalities if they focus only on onsite spillover effects while overlooking the remote spillover effects or ignoring spillovers to other firms. It is worth exploring whether firms' return-to-office policies bring the economy closer to the social optimum. I leave the extension to cross-city mobility, cross-region remote productivity spillover, and other policy considerations for future work.

Appendix

A Empirical Facts

A.1 The WFH Shock

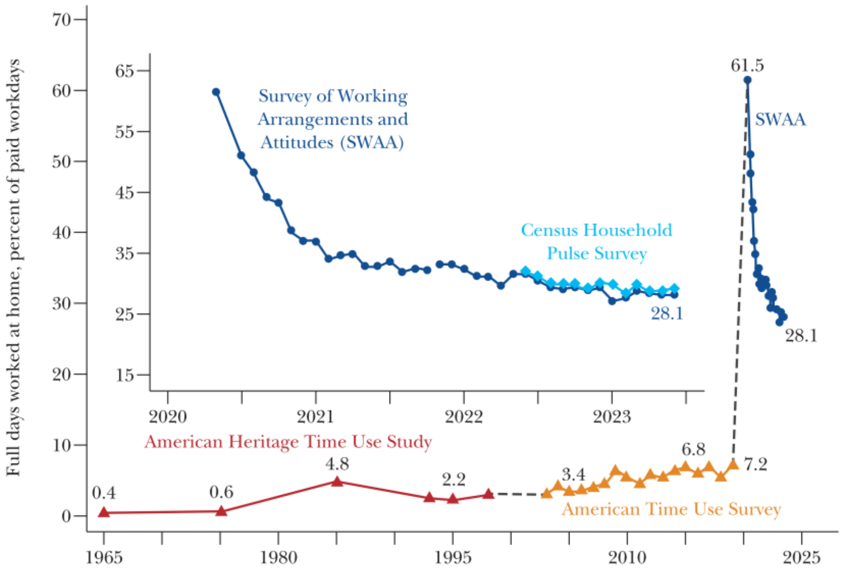


Figure 16: Historical and current WFH trend

Figure source: Barrero et al. (2023).

Table 4: Summary statics for the individual level share of time spent WFH

	Obs	Mean	Std.Dev.	Min	Max
share of time WFH (CPS 1997,2001,2004)	155,742	0.02	0.12	0.00	1.00
share of time WFH (CPS 2022,2023,2024)	1,036,136	0.14	0.31	0.00	1.00
share of time WFH (NLSY79 1988-2018)	75,298	0.04	0.14	0.00	1.00
share of time WFH (NLSY79 2020)	3,971	0.19	0.36	0.00	1.00

Note: The share of time spent WFH is measured by the ratio of hours worked from home to total work hours in a week. The sample is laborforce excluding self-employed jobs, military jobs, or unpaid jobs in family business or farming. The data sources are the CPS and NLSY79. The CPS data cover three months from the work schedule supplement in 1997, 2001, and 2004 and 20 months from the basic month survey spanning October 2022 to May 2024. The NLSY79 data include annual records from 1988 to 1993 and biennial records from 1994 to 2020. The 1988-2018 NLSY79 dataset spans 19 years. In CPS 2022,2023,2024, only workers with one job are included. For other periods of CPS and NLSY79 data, the share of time spent WFH is calculated for individuals' main jobs.

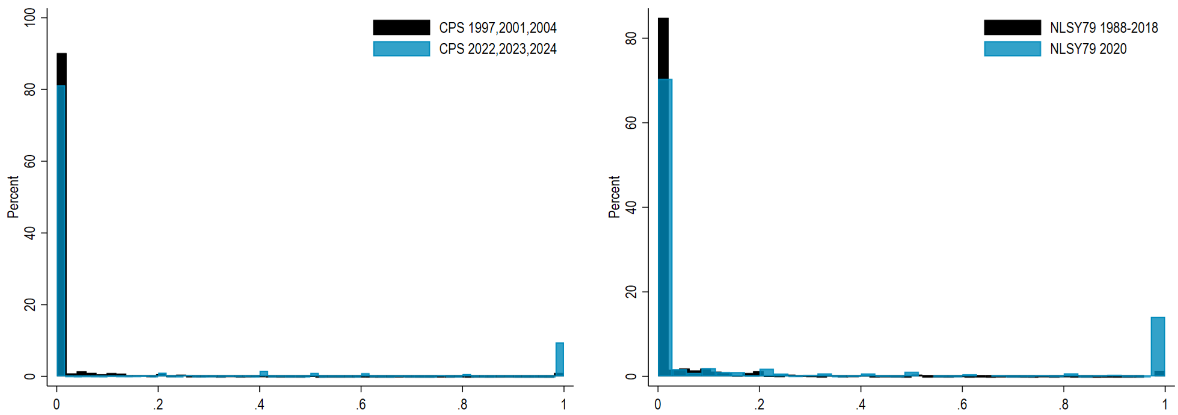


Figure 17: Distribution of share of time spent WFH at individual level

A.2 Geographic and Sectoral Mobility Patterns

Table 5: Migration Shares of Home-Based and On-Site Workers Within 1 Year

Panel A: Shares in laborforce

	2018-2019		2021-2023	
Panel A: shares in labforce	on-site worker	home-based worker	on-site worker	home-based worker
same house (%)	85.06	84.71	86.52	84.16
moved within CBSA (%)	10.25	8.15	8.74	9.33
moved between CBSAs (%)	4.27	6.44	4.26	6.04
abroad 1 year ago (%)	0.42	0.70	0.49	0.47
N (unweighted)	2,658,947		4,052,222	

Panel B: Shares in migrator

Panel B:shares in migrator	on-site worker	home-based worker	on-site worker	home-based worker
moved within CBSA (%)	68.62	53.32	64.83	58.94
moved between CBSAs (%)	28.55	42.13	31.57	38.12
abroad 1 year ago (%)	2.83	4.55	3.60	2.94
N (unweighted)	365,680		505,624	

Note: Data source is ACS. Samples are wage workers excluding unpaid family workers and military workers. Each observation's current or one year ago residence (PUMA or MIGPUMA) is matched to one CBSA with the largest population share.

Table 6: Changes in the Sector of Primary Job Within 1 Year

	2018-2019			2021-2023		
	fully WFH	hybrid	fully onsite	fully WFH	hybrid	fully onsite
no change in industry (%)	96.64	31.11	94.40	95.71	33.44	93.88
change industry 1 time (%)	3.19	64.67	5.25	4.15	61.42	5.81
change industry 2-4 times (%)	0.17	4.22	0.36	0.14	5.14	0.30
N (unweighted)	666,746			799,997		

Note: Data source is Survey of Income and Program Participation (SIPP). The survey collects observations at the person-month level. The primary job is defined as the job with the highest number of weekly working hours. Changes in occupations and industries are measured at the 2-digit SOC occupation level or 2-digit NAICS industry level. "Hybrid" refers to individuals who worked from home at least one day per week during the interview year.

B First-Order Condition for the Share of Time Spent Working Onsite in Market Equilibrium

This section shows the process of deriving the equation (7). The first-order condition of the indirect utility function with respect to θ, given unit wage w and prices P, q is:

$$\frac{w\frac{\partial \ell(\theta)}{\partial \theta}d(\theta)P^{\alpha}q^{1-\alpha} - w\ell(\theta)\frac{\partial d(\theta)}{\partial \theta}P^{\alpha}q^{1-\alpha}}{(d(\theta)P^{\alpha}q^{1-\alpha})^2} = 0$$

Thus, $\dfrac{\partial \ell(\theta)}{\partial \theta}d(\theta) = \ell(\theta)\dfrac{\partial d(\theta)}{\partial \theta}$

$$\frac{\partial \ell(\theta)/\partial \theta}{\ell(\theta)}d(\theta) = \frac{\partial d(\theta)}{\partial \theta}$$

$$\frac{\rho}{\rho-1}\left[(A(1-\theta))^{\frac{\rho-1}{\rho}} + (B\theta)^{\frac{\rho-1}{\rho}}\right]^{-1}\left[\frac{\rho}{\rho-1}(A(1-\theta))^{-\frac{1}{\rho}}(-A) + \frac{\rho}{\rho-1}(B\theta)^{-\frac{1}{\rho}}B\right]d(\theta) = \frac{\partial d(\theta)}{\partial \theta}$$

$$\left(-A^{\frac{\rho-1}{\rho}}(1-\theta)^{-\frac{1}{\rho}} + B^{\frac{\rho-1}{\rho}}\theta^{-\frac{1}{\rho}}\right)(\theta\zeta^o + (1-\theta)\zeta) = (\zeta^o - \zeta)\left((A(1-\theta))^{\frac{\rho-1}{\rho}} + (B\theta)^{\frac{\rho-1}{\rho}}\right)$$

$$B^{\frac{\rho-1}{\rho}}\theta^{-\frac{1}{\rho}}(\theta\zeta^o + (1-\theta)\zeta - (\zeta^o-\zeta)\theta) = A^{\frac{\rho-1}{\rho}}(1-\theta)^{-\frac{1}{\rho}}((\zeta^o-\zeta)(1-\theta) + \theta\zeta^o + (1-\theta)\zeta)$$

$$\frac{\theta}{1-\theta} = \left(\frac{B}{A}\right)^{\rho-1}\left(\frac{\zeta}{\zeta^o}\right)^{\rho}$$

$$\theta = \frac{1}{1 + (\frac{\zeta^o}{\zeta})^{\rho}(\frac{A}{B})^{\rho-1}}.$$

C Optimal Share of Time Spent Working Onsite in the Basic Model

In the one city one sector model with homogenous workers, the social planner's problem is

$$\max_{c,h,K,\theta} u\bar{L},$$
$$\text{s.t.} \begin{cases} u = \frac{c^{\alpha}h^{1-\alpha}}{\bar{\phi}d(\theta)}, \\ c\bar{L} + K = Y = \ell(\theta)\bar{L} \text{ (goods market clearing)}, \\ h\bar{L} = K^{\gamma}\bar{H}^{1-\gamma} \text{ (housing market clearing)}. \end{cases} \tag{C.1}$$

where $\bar{\phi} = \alpha^{\alpha}(1-\alpha)^{1-\alpha}$, $d(\theta) = \zeta^o\theta + \zeta(1-\theta)$. $\ell(\theta) = \left[(A(\theta)(1-\theta))^{\frac{\rho-1}{\rho}} + (B(\theta)\theta)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}$, $A(\theta) = \bar{a}((1-\theta)\bar{L} + \tau\theta\bar{L})^{\lambda^R}$, $B(\theta) = \bar{b}(\theta\bar{L} + \tau(1-\theta)\bar{L})^{\lambda}$.

The Lagrangian function is:

$$\mathscr{L} = u\bar{L} - \widetilde{W}\left(u - \frac{c^{\alpha}h^{1-\alpha}}{\bar{\phi}d(\theta)}\right) - \widetilde{P}\left(c\bar{L} + K - \ell(\theta)\bar{L}\right) - \widetilde{q}\left(\hbar\bar{L} - K^{r}\bar{H}^{1-r}\right). \tag{C.2}$$

The first-order conditions are:

$$\partial c : \widetilde{W}\frac{\partial u}{\partial c} = \widetilde{P}\bar{L}; \tag{C.3}$$

$$\partial h : \widetilde{W}\frac{\partial u}{\partial h} = \widetilde{q}\bar{L}; \tag{C.4}$$

$$\partial K : \widetilde{P} = \widetilde{q}\gamma K^{\gamma-1}\bar{H}^{1-\gamma}; \tag{C.5}$$

$$\partial \theta : \widetilde{W}\frac{\partial u}{\partial \theta} = -\widetilde{P}\frac{\partial \ell(\theta)}{\partial \theta}\bar{L}. \tag{C.6}$$

To solve for the socially optimal share of time spent working onsite, I first express h, K as a function of c and then express c as a function of θ. Combining first-order condition with respect to consumption and housing gives

$$h = \frac{1-\alpha}{\alpha}\frac{\widetilde{P}}{\widetilde{q}}c. \tag{C.7}$$

Combining first-order condition with respect to tradable goods used to produce housing (C.5), housing market clearing condition, and equation (C.7) yields

$$K = \frac{1-\alpha}{\alpha}\gamma c\bar{L}. \tag{C.8}$$

Substituting equation (C.8) to the goods market clearing condition yields

$$c = \frac{\alpha}{\alpha + \gamma(1-\alpha)}\ell(\theta). \tag{C.9}$$

Substituting equations (C.8) and (C.9) to equation (C.5) gives the price ratio:

$$\frac{\widetilde{P}}{\widetilde{q}} = \gamma^{\gamma}\left[\frac{1-\alpha}{\alpha + \gamma(1-\alpha)}\right]^{\gamma-1}\left(\frac{Y}{\bar{H}}\right)^{\gamma-1}. \tag{C.10}$$

Total welfare $u\bar{L} = \frac{c^{\alpha}h^{1-\alpha}}{\bar{\phi}d(\theta)}\bar{L}$ can be expressed as a function of share of time spent working onsite, where consumption and housing are determined by equations (C.7), (C.9) and (C.10).

The next step is to combine the first-order conditions to consumption (C.3), share of time spent working onsite (C.6), and equation (C.9) to obtain an equation that pins down the solution for θ. The first-order conditions to consumption (C.3) gives

$$\frac{\widetilde{W}}{\widetilde{P}} = \frac{c\bar{L}}{\alpha u} \tag{C.11}$$

The first-order conditions to share of time spent working onsite (C.6) gives

$$\frac{\widetilde{W}u}{\widetilde{P}} \frac{\partial d(\theta)/\partial\theta}{d(\theta)} = \frac{\partial \ell(\theta)}{\partial\theta}\bar{L} \tag{C.12}$$

Substituting equation (C.11) and (C.9) to equation (C.12) yields

$$\frac{\partial d(\theta)/\partial\theta}{d(\theta)} = \frac{\partial \ell(\theta)/\partial\theta}{\ell(\theta)}\tilde{\gamma}, \quad \tilde{\gamma} = \alpha + \gamma(1-\alpha). \tag{C.13}$$

Finally, I simplify the equation (C.13) to obtain an equation comparable to the one used to solve θ in market equilibrium. The social planner considers the onsite and remote productivity $(B(\theta), A(\theta))$ as a function of the share of time spent working onsite. Therefore, the percentage change in efficiency labor unit with respect to the share of time spent working onsite is:

$$\frac{\partial \ell(\theta)/\partial\theta}{\ell(\theta)} = \left[(A(\theta)(1-\theta))^{\frac{\rho-1}{\rho}} + (B(\theta)\theta)^{\frac{\rho-1}{\rho}}\right]^{-1} \tag{C.14}$$

$$\left[-(A(\theta))^{\frac{\rho-1}{\rho}}(1-\theta)^{-\frac{1}{\rho}}\left(\delta^R+1\right) + (B(\theta))^{\frac{\rho-1}{\rho}}\theta^{-\frac{1}{\rho}}(\delta+1)\right], \tag{C.15}$$

where $\delta = \frac{\partial B(\theta)/B(\theta)}{\partial\theta/\theta} = \lambda(1-\tau)\frac{\theta\bar{L}}{\theta\bar{L}+\tau(1-\theta)\bar{L}}$ denotes the elasticity of onsite productivity with respect to share of time spent working onsite. $\delta^R = \frac{\partial A(\theta)/A(\theta)}{\partial(1-\theta)/(1-\theta)} = \lambda^R(1-\tau)\frac{(1-\theta)\bar{L}}{(1-\theta)\bar{L}+\tau\theta\bar{L}}$ denotes the elasticity of remote productivity with respect the share of time spent WFH.

Substituting equation (C.15) to equation (C.13) gives

$$A(\theta)^{\frac{\rho-1}{\rho}}(1-\theta)^{\frac{-1}{\rho}}\left(\tilde{\gamma}d(\theta)(\delta^R+1)+\frac{\partial d(\theta)}{\partial\theta}(1-\theta)\right)=(B(\theta))^{\frac{\rho-1}{\rho}}\theta^{-\frac{1}{\rho}}\left(\tilde{\gamma}d(\theta)(\delta+1)-\frac{\partial d(\theta)}{\partial\theta}\theta\right). \tag{C.16}$$

Rearranging this equation yields

$$\frac{\theta}{1-\theta}=\left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1}\left[\frac{\tilde{\gamma}d(\theta)(\delta+1)-\frac{\partial d(\theta)}{\partial\theta}\theta}{\tilde{\gamma}d(\theta)(\delta^R+1)+\frac{\partial d(\theta)}{\partial\theta}(1-\theta)}\right]^{\rho} \tag{C.17}$$

$$\frac{\theta}{1-\theta}=\left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1}\left[\frac{\tilde{\gamma}d(\theta)(\delta+1)-d(\theta)+d(\theta)-\frac{\partial d(\theta)}{\partial\theta}\theta}{\tilde{\gamma}d(\theta)(\delta^R+1)-d(\theta)+d(\theta)+\frac{\partial d(\theta)}{\partial\theta}(1-\theta)}\right]^{\rho} \tag{C.18}$$

$$\frac{\theta}{1-\theta}=\left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1}\left[\frac{\zeta+d(\theta)\left(\tilde{\gamma}\left(\delta+1\right)-1\right)}{\zeta^o+d(\theta)\left(\tilde{\gamma}\left(\delta^R+1\right)-1\right)}\right]^{\rho}, \tag{C.19}$$

where $\tilde{\gamma}=\alpha+\gamma(1-\alpha)$, $A(\theta)=\bar{a}((1-\theta)\bar{L}+\tau\theta\bar{L})^{\lambda^R}$, $B(\theta)=\bar{b}(\theta\bar{L}+\tau(1-\theta)\bar{L})^{\lambda}$. $\delta=\frac{\partial B(\theta)/B(\theta)}{\partial\theta/\theta}=\lambda(1-\tau)\frac{\theta\bar{L}}{\theta\bar{L}+\tau(1-\theta)\bar{L}}$. $\delta^R=\frac{\partial A(\theta)/A(\theta)}{\partial(1-\theta)/(1-\theta)}=\lambda^R(1-\tau)\frac{(1-\theta)\bar{L}}{(1-\theta)\bar{L}+\tau\theta\bar{L}}$. $d(\theta)=\theta\zeta^o+(1-\theta)\zeta$. Another way to simplify equation (C.17) is:

$$\frac{\theta}{1-\theta}=\left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1}\left(\frac{\tilde{\gamma}(\delta+1)-\frac{\partial d(\theta)/d(\theta)}{\partial\theta/\theta}}{\tilde{\gamma}(\delta^R+1)+\frac{\partial d(\theta)/d(\theta)}{\partial\theta/(1-\theta)}}\right)^{\rho}, \tag{C.20}$$

$$=\left(\frac{B(\theta)}{A(\theta)}\right)^{\rho-1}\left(\frac{\tilde{\gamma}(\delta+1)-\delta_d}{\tilde{\gamma}(\delta^R+1)-\delta_d^R}\right)^{\rho}, \tag{C.21}$$

where $\delta_d=\frac{\partial d(\theta)/d(\theta)}{\partial\theta/\theta}=\frac{(\zeta^o-\zeta)\theta}{d(\theta)}$, $\delta_d^R=\frac{\partial d(\theta)/d(\theta)}{\partial(1-\theta)/(1-\theta)}=\frac{(\zeta-\zeta^o)(1-\theta)}{d(\theta)}$.

D Optimal Policies in the Basic Model

D.1 Using Income Tax to Subsidize Onsite Work or Remote Work

If the social planner uses income tax to subsidize onsite work or remote work, The disposable income under this policy becomes

$$\mathscr{W}(\theta)=I(\theta)h(\theta),\quad h(\theta)=(1-\mathscr{T})\,(1+s\theta)\,, \tag{D.1}$$

where $\mathscr{T}$ is income tax rate. s is the onsite subsidy (tax) rate. $s > 0$ corresponds to subsidizing onsite work or taxing remote work; $s < 0$ corresponds to taxing onsite work or subsidizing remote work.

The social planner's budget constraint is

$$I(\theta)\mathscr{T}\bar{L} = I(\theta)(1-\mathscr{T})s\theta\bar{L}. \tag{D.2}$$

Simplifying this equation gives the relationship between income tax and subsidy: $\mathscr{T} = \frac{s\theta}{1+s\theta}$. Therefore, the share of onsite subsidy (tax) in gross income is:

$$SubsidySh_{inctax} = \frac{I(\theta)(1-\mathscr{T})s\theta}{I(\theta)} = \frac{s\theta}{1+s\theta}. \tag{D.3}$$

This share equals the share of income tax in gross income $\frac{I(\theta)(1-\mathscr{T})}{I(\theta)} = 1-\mathscr{T}$.

Under the policy, the first-order condition to θ in market equilibrium becomes

$$\frac{\partial d(\theta)/\partial\theta}{d(\theta)} = \frac{\partial \ell(\theta)/\partial\theta}{\ell(\theta)} + \frac{\partial h(\theta)/\partial\theta}{h(\theta)}, \tag{D.4}$$

where $\partial\ell(\theta)$ differs from the social planner problem in that onsite and remote productivity (B, A) are taken as constant rather than a function of θ.

Workers' choice after policy is

$$\frac{\theta}{1-\theta} = \left(\frac{B}{A}\right)^{\rho-1}\left(\frac{\zeta + H(\theta)d(\theta)\theta}{\zeta^o - H(\theta)d(\theta)(1-\theta)}\right)^{\rho}, \tag{D.5}$$

where $H(\theta) = \frac{dh(\theta)/d\theta}{h(\theta)} = \frac{s}{(1+s\theta)}$. Equating post-policy workers' choice (D.5) to the social optimum equation (11) gives the solution for $H(\theta)$ and the share of subsidy in gross income:

$$H(\theta) = \frac{\widetilde{Z}\zeta^o - \zeta}{d(\theta)\theta + \widetilde{Z}d(\theta)(1-\theta)}, \quad SubsidySh_{inctax} = H(\theta)\theta, \tag{D.6}$$

$$s = \frac{H(\theta)}{1-H(\theta)\theta} \tag{D.7}$$

where $\widetilde{Z} = \frac{\zeta + d(\theta)(\widetilde{\gamma}(\delta+1)-1)}{\zeta^o + d(\theta)(\widetilde{\gamma}(\delta^R+1)-1)}$.

D.2 Taxing (Subsidizing) Remote Work and Subsidizing (Taxing) Onsite Work

An alternative policy is that tax and subsidy are all related to the sources of externalities. If the social planner taxes (subsidies) remote work and subsidies (taxes) the onsite work, the disposable income becomes

$$\mathscr{W}(\theta) = I(\theta)g(\theta), \quad g(\theta) = \left(1 - \mathscr{t}(1-\theta)\right)(1 + t\theta)\,, \tag{D.8}$$

where $\mathscr{t}$ is the tax (subsidy) rate based on share of time spent WFH, t is subsidy (tax) rate based on share of time spent working onsite. The social planner's budget constraint is

$$I(\theta)\mathscr{t}(1-\theta)\bar{L} = I(\theta)(1 - \mathscr{t}(1-\theta))t\theta\bar{L}. \tag{D.9}$$

Simplifying this equation gives the relationship between the tax and the subsidy:

$$\mathscr{t}(1-\theta) = \frac{t\theta}{1+t\theta}. \tag{D.10}$$

Thus, the share of onsite subsidy (tax) in gross income can be derived as

$$SubsidySh_{exttax} = \frac{I(\theta)(1 - \mathscr{t}(1-\theta))t\theta}{I(\theta)} = \frac{t\theta}{1+t\theta}. \tag{D.11}$$

This share equals the share of remote tax (subsidy) in gross income $\frac{I(\theta)\mathscr{t}(1-\theta)}{I(\theta)} = \mathscr{t}(1-\theta)$.

Workers' choice after policy is

$$\frac{\theta}{1-\theta} = \left(\frac{B}{A}\right)^{\rho-1}\left(\frac{\zeta + G(\theta)d(\theta)\theta}{\zeta^o - G(\theta)d(\theta)(1-\theta)}\right)^{\rho}, \tag{D.12}$$

where $G(\theta) = \frac{\partial g(\theta)/\partial\theta}{g(\theta)} = \frac{t\left(t\theta^2+1\right)}{(1-\theta)(1+t\theta)}$. Let post-policy workers' choice (D.12) equals the social optimum equation (11), I obtain

$$\theta^2 t^2 + \left(1 - \theta(1-\theta)G^*\right)t - G^*(1-\theta) = 0, \tag{D.13}$$

Therefore, the solution for the onsite subsidy (tax) rate is as follows:

$$t = \frac{1}{2\theta^2}\left[-1+\theta(1-\theta)G^* + \sqrt{(1-\theta(1-\theta)G^*)^2 + 4\theta^2 G^*(1-\theta)}\right], \tag{D.14}$$

where $G^* = \frac{\widetilde{Z}\zeta^o - \zeta}{d(\theta)\theta + \widetilde{Z}d(\theta)(1-\theta)}$, $\widetilde{Z} = \frac{\zeta + d(\theta)(\widetilde{\gamma}(\delta+1)-1)}{\zeta^o + d(\theta)(\widetilde{\gamma}(\delta^R+1)-1)}$.

E Social Planner's Problem in the Full Model

In the full model, workers in the same city have the same expected utility. The social planner maximizes the utility of an arbitrary city i subject to the utility frontier defined by resource constraints of all cities:

$$\max \sum_i V_i \bar{L}_i$$

subject to

1. labor mobility constraint:

$$V_i \leq v_{is}^m \left(\frac{\bar{L}_i}{L_{is}^m}\right)^{\frac{1}{\sigma}} \ \forall i, s, m, \tag{E.1}$$

where $v_{is}^m = \frac{\left(c_{is}^m\right)^\alpha \left(h_{is}^m\right)^{1-\alpha} \varepsilon_{is}^m}{\bar{\phi} d_{is}^m\left(\theta_{is}^m\right)}$, $\bar{\phi} = \alpha^\alpha(1-\alpha)$, $d_{is}^m(\theta_{is}^m) = e^{kt_i}\theta_{is}^m + \zeta_{is}(1-\theta_{is}^m)$, $\theta_{is}^o = 1, \theta_{is}^h = \theta_{is} \in (0,1)$, and $\theta_{is}^f = 0$;

2. tradable goods constraint:

$$\sum_n Q_{in} \leq Y_i \quad \forall i, \tag{E.2}$$

$$\sum_s \sum_m c_{is}^m L_{is}^m + K_i \leq Q_i, \quad Q_i = \sum_n Q_{ni} \quad \forall i, \tag{E.3}$$

where

$$Y_i = \left[\sum_s \bar{y}_{is} \left(y_{is} \right)^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}, \qquad y_{is} = B_{is} L_{is}^m + \beta_{is} \ell_{is} \left(\theta_{is} \right) L_{is}^h + A_{is} L_{is}^f,$$

$$\ell_{is}(\theta_{is}) = [(A_{is}(\theta_{is})(1-\theta_{is}))^{\frac{\rho-1}{\rho}} + (B_{is}(\theta_{is})\theta_{is})^{\frac{\rho-1}{\rho}}]^{\frac{\rho}{\rho-1}},$$

$$B_{is}(\theta_{is}) = \bar{b}_{is} \left[L_{is}^B(\theta_{is}) + \tau L_{is}^A(\theta_{is}) \right]^{\lambda}, A_{is}(\theta_{is}) = \bar{a}_{is} [L_{is}^A(\theta_{is}) + \tau L_{is}^B(\theta_{is})]^{\lambda^R},$$

$$L_{is}^B(\theta_{is}) = L_{is}^o + \theta_{is} L_{is}^h, \qquad L_{is}^A(\theta_{is}) = L_{is}^f + (1-\theta_{is}) L_{is}^h;$$

Q_{ni} refers to the tradable goods produced in city n and sold in city i. $\sum_n Q_{in}$ is the total tradable goods produced in city i and sold to other cities (total export). $Q_i = \sum_n Q_{ni}$ is the total tradable goods that city i purchased from other cities (total import).

3. housing market clearing:

$$\sum_s \sum_m \hbar_{is}^m L_{is}^m \le K_i^r \bar{H}_i^{1-r} \quad \forall i; \tag{E.4}$$

4. labor market clearing:

$$\sum_s \sum_m L_{is}^m = \bar{L}_i \quad \forall i. \tag{E.5}$$

The Lagrangian function for the social planner problem is:

$$\begin{aligned} \mathscr{L} = & \sum_i V_i \bar{L}_i - \sum_i \sum_s \widetilde{W}_{is}^m \left(V_i - v_{is}^m \left(\frac{\bar{L}_i}{L_{is}^m} \right)^{\frac{1}{\sigma}} \right) \\ & - \sum_i \widetilde{p}_i \left(\sum_n Q_{in} - Y_i \right) - \sum_i \widetilde{P}_i \left(\sum_s \sum_m c_{is}^m L_{is}^m + K_i - Q_i \right) \\ & - \sum_i \widetilde{q}_i \left(\sum_s \sum_m \hbar_{is}^m L_{is}^m - K_i^r \bar{H}_i^{1-r} \right) - \sum_i e_i \left(\sum_s \sum_m L_{is}^m - \bar{L}_i \right). \end{aligned} \tag{E.6}$$

Under the free trade assumption, prices of tradable goods are equalized across cities. Final goods price corresponds to the multiplier of goods market clearing. So, I use an equalized multiplier $\widetilde{P}_i = \widetilde{P}$ to match with the price vector in the market equilibrium.

The first-order conditions with respect to consumption, housing, share of time spent working onsite, tradable goods used in housing production, and employment are:

$$\partial c_{is}^m : \widetilde{W}_{is}^m \frac{\partial v_{is}^m}{\partial c_{is}^m} \left(\frac{\bar{L}_i}{L_{is}^m} \right)^{\frac{1}{\sigma}} = \widetilde{P}_i L_{is}^m, \forall i, s, m; \tag{E.7}$$

$$\partial h_{is}^m : \widetilde{W}_{is}^m \frac{\partial v_{is}^m}{\partial h_{is}^m} \left(\frac{\bar{L}_i}{L_{is}^m} \right)^{\frac{1}{\sigma}} = \widetilde{q}_i L_{is}^m, \forall i, s, m; \tag{E.8}$$

$$\partial K_i : \widetilde{P}_i = \tilde{q}_i \gamma K_i^{\gamma-1} \bar{H}_i^{1-\gamma}, \forall i; \tag{E.9}$$

$$\partial \theta_{is} : \widetilde{W}_{is}^h \frac{\partial v_{is}^h}{\partial \theta_{is}} \left(\frac{\bar{L}_i}{L_{is}^h} \right)^{\frac{1}{\sigma}} = -\widetilde{p}_i \frac{\partial Y_i}{\partial \theta_{is}}, \forall i, s; \tag{E.10}$$

$$\partial Q_{ni} : \widetilde{p}_n = \widetilde{P}_i, \forall i, n; \tag{E.11}$$

$$\partial L_{is}^m : -\widetilde{W}_{is}^m v_{is}^m \frac{1}{\sigma} \left(\frac{\bar{L}_i}{L_{is}^m} \right)^{\frac{1}{\sigma}} \frac{1}{L_{is}^m} + \widetilde{p}_i \frac{\partial Y_i}{\partial L_{is}^m} - \widetilde{P}_i c_{is}^m - \tilde{q}_i h_{is}^m - e_i = 0, \forall i, s, m. \tag{E.12}$$

According to the first-order condition with respect to trade flow (E.11), the multiplier $\widetilde{p}_n = \widetilde{P}_i = \widetilde{P}$.

Consumption and Housing

Substituting labor mobility constraint (E.1) to first-order conditions with respect to consumption (E.7) and housing (E.8) yields

$$c_{is}^m = \frac{\alpha \widetilde{W}_{is}^m V_i}{\widetilde{P} L_{is}^m}, \quad h_{is}^m = \frac{(1-\alpha) \widetilde{W}_{is}^m V_i}{\widetilde{q}_i L_{is}^m}. \tag{E.13}$$

Define $\widetilde{I}_{is}^m = \widetilde{P} c_{is}^m + \widetilde{q}_i h_{is}^m$. Based on equation (E.13), $\widetilde{I}_{is}^m$ becomes

$$\widetilde{I}_{is}^m = \widetilde{P} c_{is}^m + \widetilde{q}_i h_{is}^m = \frac{\widetilde{W}_{is}^m V_i}{L_{is}^m}. \tag{E.14}$$

Then, the solution for consumption and housing can be written as

$$c_{is}^m = \frac{\alpha \widetilde{I}_{is}^m}{\widetilde{P}}, \tag{E.15}$$

$$h_{is}^m = \frac{(1-\alpha) \widetilde{I}_{is}^m}{\widetilde{q}_i}. \tag{E.16}$$

Therefore, utility can be expressed as $v_{is}^m = \frac{\widetilde{I}_{is}^m \epsilon_{is}^m}{\widetilde{P}^\alpha \widetilde{q}_i^{\,1-\alpha} d_{is}^m(\theta_{is}^m)}$.

Final goods used to produce housing

Combining the first-order condition with respect to K_i (E.9) and housing market clearing (E.8) yields

$$\sum_s \sum_m h_{is}^m L_{is}^m = \frac{K_i}{\gamma} \frac{\widetilde{P}}{\widetilde{q}_i}. \tag{E.17}$$

Substituting $h_{is}^m = \frac{(1-\alpha)\widetilde{I}_{is}}{\widetilde{q}_i}$ to equation (E.17) yields

$$K_i = \frac{\gamma(1-\alpha)}{\widetilde{P}} \sum_s \sum_m \widetilde{I}_{is}^m L_{is}^m. \tag{E.18}$$

Based on equation (E.9), the multiplier of the housing market clearing constraint is:

$$\widetilde{q}_i = \frac{\widetilde{P}}{\gamma K_i^{\gamma-1} \bar{H}_i^{1-\gamma}}. \tag{E.19}$$

Substituting equation (E.18) to this equation gives the solution for $\widetilde{q}_i$:

$$\widetilde{q}_i = \left(\frac{\widetilde{P}}{\gamma}\right)^{\gamma} \left(\frac{(1-\alpha)\sum_s \sum_m \widetilde{I}_{is}^m L_{is}^m}{\bar{H}_i}\right)^{1-\gamma}. \tag{E.20}$$

Employment

Rearranging labor mobility constraint (E.1) gives $L_{is}^m = \left(\frac{v_{is}^m}{V_i}\right)^{\sigma} \bar{L}_i$. Substituting this equation into the labor market clearing yields $V_i = \left(\sum_s \sum_m (v_{is}^m)^{\sigma}\right)^{\frac{1}{\sigma}}$. The solution for employment is $L_{is}^m = \left(\frac{v_{is}^m}{V_i}\right)^{\sigma} \bar{L}_i$, where $v_{is}^m = \frac{\widetilde{I}_{is}^m \epsilon_{is}^m}{\widetilde{P}^{\alpha} \widetilde{q}_i^{1-\alpha} d_{is}^m(\theta_{is}^m)}$, $V_i = \left(\sum_s \sum_m (v_{is}^m)^{\sigma}\right)^{\frac{1}{\sigma}}$. Simplifying this equation gives the solution for employment:

$$L_{is}^{m*} = \left(\frac{\widetilde{\phi}_{is}^m}{\widetilde{\widetilde{\Phi}}_i}\right)^{\sigma} \bar{L}_i, \tag{E.21}$$

where $\widetilde{\phi}_{is}^m = \frac{\widetilde{I}_{is}^m \epsilon_{is}^m}{d_{is}^m(\theta_{is}^m)}$, $\widetilde{\widetilde{\Phi}}_i = \left(\sum_s \sum_m \left(\widetilde{\phi}_{is}^m\right)^{\sigma}\right)^{\frac{1}{\sigma}}$.

Individual Expenditure

Substituting labor mobility constraint (E.1) and equation (E.14) to the first-order condition with respect to employment (E.12) yields

$$\widetilde{I}_{is}^{m} = \frac{\sigma}{\sigma+1}\left(\widetilde{P}\frac{\partial Y_i}{\partial L_{is}^{m}} - e_i\right), \tag{E.22}$$

where $\frac{\partial Y_i}{\partial L_{is}^{m}} = \frac{\partial Y_i}{\partial y_{is}}\left(y_{is}^{m} + \sum_{m'} \frac{\partial y_{is}^{m'}}{\partial L_{is}^{m}}\right)$. Substituting equation (E.15) and (E.18) to the import constraint (E.3) gives:

$$Q_i = \frac{\alpha+\gamma(1-\alpha)}{\widetilde{P}}\sum_s\sum_m \widetilde{I}_{is}^{m} L_{is}^{m}. \tag{E.23}$$

Substituting equation (E.22) to equation (E.23) gives the solution for the multiplier of labor market clearing:

$$e_i = \widetilde{P}\sum_s\sum_m\left(\frac{L_{is}^{m}}{\bar{L}_i}\frac{\partial Y_i}{\partial L_{is}^{m}}\right) - \frac{\sigma+1}{\sigma}\sum_s\sum_m \frac{L_{is}^{m}}{\bar{L}_i}\widetilde{I}_{is}^{m}, \tag{E.24}$$

Substituting equation (E.24) to equation (E.22) gives the solution for a worker's expenditure:

$$\widetilde{I}_{is}^{m} = \frac{\sigma\widetilde{P}}{\sigma+1}\frac{\partial Y_i}{\partial L_{is}^{m}} - \frac{\sigma\widetilde{P}}{\sigma+1}\sum_s\sum_m\left(\frac{L_{is}^{m}}{\bar{L}_i}\frac{\partial Y_i}{\partial L_{is}^{m}}\right) + \sum_s\sum_m \frac{L_{is}^{m}}{\bar{L}_i}\widetilde{I}_{is}^{m}, \tag{E.25}$$

$$\text{where } \frac{\partial Y_i}{\partial L_{is}^{m}} = \bar{y}_{is}\left(\frac{Y_i}{y_{is}}\right)^{\frac{1}{\eta}}\frac{\partial y_{is}}{\partial L_{is}^{m}} \quad \forall m\in\{o,h,f\}; \tag{E.26}$$

$$\frac{\partial y_{is}}{\partial L_{is}^{m}} = \frac{\partial B_{is}}{\partial L_{is}^{m}}L_{is}^{o} + \beta_{is}L_{is}^{h}\left(\frac{\partial \ell_{is}(\theta_{is})}{\partial A_{is}}\frac{\partial A_{is}}{\partial L_{is}^{m}} + \frac{\partial \ell_{is}(\theta_{is})}{\partial B_{is}}\frac{\partial B_{is}}{\partial L_{is}^{m}}\right) + \frac{\partial A_{is}}{\partial L_{is}^{m}}L_{is}^{f} + y_{is}^{m} \quad \forall m; \tag{E.27}$$

$$y_{is}^{o} = B_{is}, \quad y_{is}^{h} = \beta_{is}\ell_{is}(\theta_{is}), \quad y_{is}^{h} = A_{is}; \tag{E.28}$$

$$\frac{\partial B_{is}}{\partial L_{is}^{o}} = \frac{\lambda B_{is}}{L_{is}^{B}+\tau L_{is}^{A}}, \quad \frac{\partial B_{is}}{\partial L_{is}^{h}} = \frac{\lambda B_{is}(\theta_{is}+\tau(1-\theta_{is}))}{L_{is}^{B}+\tau L_{is}^{A}}, \quad \frac{\partial B_{is}}{\partial L_{is}^{f}} = \frac{\tau\lambda B_{is}}{L_{is}^{B}+\tau L_{is}^{A}}; \tag{E.29}$$

$$\frac{\partial A_{is}}{\partial L_{is}^{o}} = \frac{\tau\lambda^{R} A_{is}}{L_{is}^{A}+\tau L_{is}^{B}}, \quad \frac{\partial A_{is}}{\partial L_{is}^{h}} = \frac{\lambda^{R} A_{is}(1-\theta_{is}+\tau\theta_{is})}{L_{is}^{A}+\tau L_{is}^{B}}, \quad \frac{\partial A_{is}}{\partial L_{is}^{f}} = \frac{\lambda^{R} A_{is}}{L_{is}^{A}+\tau L_{is}^{B}}; \tag{E.30}$$

$$\frac{\partial \ell_{is}(\theta_{is})}{\partial B_{is}} = \left(\frac{\ell_{is}(\theta_{is})}{B_{is}}\right)^{\frac{1}{\rho}} (\theta_{is})^{\frac{\rho-1}{\rho}}, \quad \frac{\partial \ell_{is}(\theta_{is})}{\partial A_{is}} = \left(\frac{\ell_{is}(\theta_{is})}{A_{is}}\right)^{\frac{1}{\rho}} (1-\theta_{is})^{\frac{\rho-1}{\rho}}. \tag{E.31}$$

Shadow Price of Tradable Goods

According to tradable goods constraints (E.3, E.2), total consumption of the economy equals total production:

$$\sum_i Q_i = \sum_i \sum_n Q_{ni} = \sum_i Y_i. \tag{E.32}$$

Substituting equation (E.23) to equation (E.32) and rearranging the equation yields the multiplier of tradable goods constraints:

$$\widetilde{P} = \frac{(\alpha + \gamma(1-\alpha)) \sum_i \sum_s \sum_m \widetilde{I}_{is}^m L_{is}^m}{\sum_i Y_i}. \tag{E.33}$$

Hybrid workers' share of time spent working onsite

Substituting labor mobility constraint (E.1) to the first-order condition with respect to share of time hybrid workers spend working onsite (E.10) gives

$$\frac{\widetilde{W}_{is}^h}{\widetilde{P}} = \frac{1}{V_i} \frac{\partial Y_i}{\partial \theta_{is}} \frac{d_{is}^h(\theta_{is})}{\partial d_{is}^h(\theta_{is})/\partial \theta_{is}}, \tag{E.34}$$

where $d_{is}^h(\theta_{is}) = e^{kt_i}\theta_{is} + \zeta_{is}(1-\theta_{is})$. Substituting equation (E.34) into $c_{is}^m = \frac{\alpha \widetilde{W}_{is}^m V_i}{\widetilde{P} L_{is}^m}$ gives the total consumption of hybrid workers in a city-sector cell:

$$c_{is}^h L_{is}^h = \alpha \frac{d_{is}^h(\theta_{is})}{\partial d_{is}^h(\theta_{is})/\partial \theta_{is}} \frac{\partial Y_i}{\partial \theta_{is}}, \tag{E.35}$$

Substituting (E.15) into equation (E.35) and rearranging the equation yields:

$$\frac{\partial d_{is}^h(\theta_{is})/\partial \theta_{is}}{d_{is}^h(\theta_{is})} = F_{is} \frac{\partial \ell_{is}(\theta_{is})/\partial \theta_{is}}{\ell_{is}(\theta_{is})}, \tag{E.36}$$

$$F_{is} = \frac{\widetilde{P} \frac{\partial Y_i}{\partial \theta_{is}}}{\widetilde{I}_{is}^h L_{is}^h \frac{\partial \ell_{is}(\theta_{is})/\partial \theta_{is}}{\ell_{is}(\theta_{is})}}, \tag{E.37}$$

where the numerator of F_{is} denotes the value of the marginal product of hybrid workers' onsite work. It measures the change of city i's total product in response to the change in share of time hybrid workers spend working onsite in sector s. The denominator of F_{is} reflects the change in the total income of all hybrid workers in the sector s in response to the change in share of time hybrid workers spend working onsite. Overall, F_{is} compares city-wide effect to the sector-work mode effect.

Similar to Appendix C, the interior solution of share of time hybrid workers spend working onsite is characterized by the following equation:

$$\frac{\theta_{is}^*}{1-\theta_{is}^*} = \left(\frac{B_{is}(\theta_{is}^*)}{A_{is}(\theta_{is}^*)}\right)^{\rho-1} \left(\frac{\zeta_{is} + d_{is}^h(\theta_{is}^*)(F_{is}(\delta_{is}+1)-1)}{e^{kt_i} + d_{is}^h(\theta_{is}^*)(F_{is}(\delta_{is}^R+1)-1)}\right)^{\rho}, \tag{E.38}$$

$$\text{where } F_{is} = \frac{\widetilde{P}\frac{\partial Y_i}{\partial \theta_{is}}}{\widetilde{I}_{is}^h L_{is}^h \frac{\partial \ell_{is}(\theta_{is})/\partial\theta_{is}}{\ell_{is}(\theta_{is})}}, \tag{E.39}$$

$$\frac{\partial Y_i}{\partial \theta_{is}} = \bar{y}_{is}\left(\frac{Y_i}{y_{is}}\right)^{\frac{1}{\eta}} \left(\frac{\partial B_{is}}{\partial \theta_{is}} L_{is}^o + \beta_{is}\frac{\ell_{is}(\theta_{is})}{\partial \theta_{is}} L_{is}^h, + \frac{\partial A_{is}}{\partial \theta_{is}} L_{is}^f\right), \tag{E.40}$$

$$\frac{\partial B_{is}}{\partial \theta_{is}} = \bar{b}_{is}\lambda(L_{is}^B + \tau L_{is}^A)^{\lambda-1}(1-\tau)L_{is}^h, \quad \frac{\partial A_{is}}{\partial \theta_{is}} = \bar{a}_{is}\lambda^R(L_{is}^A + \tau L_{is}^B)^{\lambda^R-1}(\tau-1)L_{is}^h, \tag{E.41}$$

$$\frac{\partial \ell_{is}}{\partial \theta_{is}} = (\ell_{is}(\theta_{is}))^{\frac{1}{\rho}} \left[-(A_{is}(\theta_{is}))^{1-\frac{1}{\rho}}(1-\theta_{is})^{-\frac{1}{\rho}}(\delta_{is}^R+1) + (B_{is}(\theta_{is}))^{1-\frac{1}{\rho}}\theta_{is}^{-\frac{1}{\rho}}(\delta_{is}+1)\right], \tag{E.42}$$

$$\widetilde{I}_{is}^h = \frac{\sigma\widetilde{P}}{\sigma+1}\frac{\partial Y_i}{\partial L_{is}^h} - \frac{\sigma}{\sigma+1}\frac{\widetilde{P}}{\bar{L}_i}\sum_s\sum_m\left(L_{is}^m\frac{\partial Y_i}{\partial L_{is}^m}\right) + \frac{1}{\bar{L}_i}\sum_s\sum_m \widetilde{I}_{is}^m L_{is}^m, \tag{E.43}$$

$$\delta_{is} = \frac{\partial B_{is}(\theta_{is})/B_{is}}{\partial\theta_{is}(\theta_{is})/\theta_{is}} = \lambda(1-\tau)\frac{\theta_{is}L_{is}^h}{L_{is}^B + \tau L_{is}^A}, \tag{E.44}$$

$$\delta_{is}^R = \frac{\partial A_{is}(\theta_{is})/A_{is}}{\partial(1-\theta_{is})/(1-\theta_{is})} = \lambda^R(1-\tau)\frac{(1-\theta_{is})L_{is}^h}{L_{is}^A + \tau L_{is}^B}, \tag{E.45}$$

$$L_{is}^B = L_{is}^o + \theta_{is}L_{is}^h, \quad L_{is}^A = L_{is}^f + (1-\theta_{is})L_{is}^h. \tag{E.46}$$

Equation (E.38) can be expressed as the following as well:

$$\frac{\theta_{is}^*}{1-\theta_{is}^*} = \left(\frac{B(\theta_{is}^*)}{A(\theta_{is}^*)}\right)^{\rho-1}\left(\frac{F_{is}(\delta_{is}+1) - \delta_{d,is}}{F_{is}(\delta_{is}^R+1) - \delta_{d,is}^R}\right)^{\rho}. \tag{E.47}$$

where $\delta_{d,is} = \frac{\partial d_{is}(\theta_{is})/d_{is}(\theta_{is})}{\partial \theta_{is}/\theta_{is}} = \frac{(e^{kt_i}-\zeta_{is})\theta_{is}}{d(\theta_{is})}, \delta^R_{d,is} = \frac{\partial d_{is}(\theta_{is})/d_{is}(\theta_{is})}{\partial(1-\theta_{is})/(1-\theta_{is})} = \frac{(\zeta_{is}-e^{kt_i})(1-\theta_{is})}{d(\theta_{is})}$.

Equations (E.15), (E.16), (F.53), (E.18), and (E.38) correspond to equations characterizing solutions for consumption, housing, employment (44), tradable goods in housing production (47), share of time hybrid workers spend working onsite (27) in the market equilibrium. $\widetilde{I}^m_{is}$, $\widetilde{P}$, $\widetilde{q}_i$ correspond to income (I^m_{is}), goods price index (P), and housing prices (q_i) in market equilibrium. The socially optimal allocation consists of consumption, housing, tradable goods used to produce housing, employment, and share of time hybrid workers spend working onsite characterized by equations (E.15), (E.16), (E.18), (F.53), (E.25), (F.49), and (E.38).

F The Algorithm for Solving Socially Optimal Allocation

Given parameters $(\sigma, \eta, \rho, k, \tau, \lambda, \lambda^R)$, exogenous variables $(\bar{L}_i, t_i)$, and shifters $(\bar{y}_{is}, \bar{b}_{is}, \bar{a}_{is}, \beta_{is}, \zeta_{is}, \epsilon^o_{is}, \epsilon^h_{is}, \epsilon^f_{is})$, I solve for socially optimal expenditure $(\widetilde{I}^o_{is}, \widetilde{I}^h_{is}, \widetilde{I}^f_{is})$, employment $(L^o_{is}, L^h_{is}, L^f_{is})$, and share of time hybrid workers spend working onsite (θ_{is}) using the following algorithm:

1. Guess L^o_{is}, L^h_{is}, and L^f_{is} such that $\sum_s (L^o_{is} + L^h_{is} + L^f_{is}) = \bar{L}_i$. Guess expenditure $\widetilde{I}^o_{is}, \widetilde{I}^h_{is}, \widetilde{I}^f_{is}$, and the share of time hybrid workers spend working onsite θ_{is}.

2. Calculate expenditure

$$\widetilde{I}^m_{is,new} = \frac{\sigma}{\sigma+1}\left(\widetilde{P}\frac{\partial Y_i}{\partial L^m_{is}} - e_i\right), \tag{F.48}$$

$$\text{where } \widetilde{P} = \frac{(\alpha+\gamma(1-\alpha))\sum_i \sum_s \sum_m \widetilde{I}^m_{is} L^m_{is}}{\sum_i Y_i}. \tag{F.49}$$

$$\frac{\partial Y_i}{\partial L^m_{is}} = \frac{\partial Y_i}{\partial y_{is}}\left(y^m_{is} + \sum_{m' \in \{o,h,f\}} \frac{\partial y^{m'}_{is}}{\partial L^m_{is}}\right), \tag{F.50}$$

$$y^o_{is} = B_{is}(L^m_{is}, \theta_{is}), y^h_{is} = \beta_{is}\ell_{is}(L^m_{is}, \theta_{is}), y^h_{is} = A_{is}(L^m_{is}, \theta_{is}). \tag{F.51}$$

$$e_i = \widetilde{P}\sum_s \sum_m \left(\frac{L^m_{is}}{\bar{L}_i}\frac{\partial Y_i}{\partial L^m_{is}}\right) - \frac{\sigma+1}{\sigma}\sum_s \sum_m \frac{L^m_{is}}{\bar{L}_i}\widetilde{I}^m_{is}, \tag{F.52}$$

3. Update employment
$$L_{is,new}^{m*} = \left(\frac{\widetilde{\phi}_{is}^m}{\widetilde{\widetilde{\Phi}}_i}\right)^{\sigma} \bar{L}_i, \tag{F.53}$$
where $\widetilde{\phi}_{is}^m = \frac{\widetilde{I}_{is}^m \epsilon_{is}^m}{d_{is}^m(\theta_{is}^m)}$, $\widetilde{\widetilde{\Phi}}_i = \left(\sum_s \sum_m \left(\widetilde{\phi}_{is}^m\right)^{\sigma}\right)^{\frac{1}{\sigma}}$.

4. Update share of time hybrid workers spend working onsite:
$$\theta_{is,new} = \left[1 + \left(\frac{B(\theta_{is})}{A(\theta_{is})}\right)^{1-\rho} \left[\frac{\zeta_{is} + d_{is}(\theta_{is})(F_{is}(\delta_{is}+1)-1)}{e^{kt_i} + d_{is}(\theta_{is})(F_{is}(\delta_{is}^R+1)-1)}\right]^{-\rho}\right]^{-1}, \tag{F.54}$$
$$\text{where } \delta_{is} = \lambda(1-\tau)\frac{\theta_{is}L_{is}^h}{L_{is}^B + \tau L_{is}^A}, \delta_{is}^R = \lambda^R(1-\tau)\frac{(1-\theta_{is})L_{is}^h}{L_{is}^A + \tau L_{is}^B}, \tag{F.55}$$
$$L_{is}^B(\theta_{is}) = L_{is}^o + \theta_{is}L_{is}^h, L_{is}^A(\theta_{is}) = L_{is}^f + (1-\theta_{is})L_{is}^h. \tag{F.56}$$
$$F_{is} = \frac{\widetilde{P}\frac{\partial Y_i}{\partial \theta_{is}}}{\widetilde{I}_{is}^h L_{is}^h \frac{\partial \ell_{is}(\theta_{is})/\partial \theta_{is}}{\ell_{is}(\theta_{is})}}. \tag{F.57}$$

5. $max\left\{\mid \frac{\theta_{is,new}}{\theta_{is}} - 1 \mid\right\}$, If $max\left\{\left|\frac{\theta_{is,new}}{\theta_{is}} - 1\right|, \left|\frac{\widetilde{I}_{is,new}^m}{\widetilde{I}_{is}^m} - 1\right|, \left|\frac{L_{is,new}^m}{L_{is}^m} - 1\right|, \forall i, s, m \in \{o, h, b\}\right\}$ is larger than a sufficient small number, set set $x = 0.5x + 0.5x_{new}, x \in \{\theta_{is}, \widetilde{I}_{is}^m, L_{is}^m\}$ and start from step 2, otherwise obtain the solution for $\theta_{is}, \widetilde{I}_{is}^m, L_{is}^m$.

G Optimal Policies in the Full Model

This section discusses policies for achieving a socially optimal share of time spent working onsite for hybrid workers similar to the basic model (section D). The implicit assumption is that there is no transfer across cities and sectors. If the social planner uses hybrid workers' income tax to subsidy (tax) hybrid workers' onsite work time, the share of onsite subsidy (tax) in gross income is:

$$SubsidySh_{is}^{inctax} = \frac{\widetilde{Z}_{is}e^{kt_i} - \zeta_{is}}{d_{is}(\theta_{is}^*)\theta_{is}^* + \widetilde{Z}_{is}d_{is}(\theta_{is}^*)(1-\theta_{is}^*)}\theta_{is}^*, \tag{G.1}$$

where $\widetilde{Z}_{is} = \frac{\zeta_{is} + d_{is}^h(\theta_{is}^*)(F_{is}(\delta_{is}+1)-1)}{e^{kt_i} + d_{is}^h(\theta_{is}^*)(F_{is}(\delta_{is}^R+1)-1)}$.

If the social planner taxes (subsidies) hybrid workers' remote work time and subsidies (taxes) the hybrid workers' onsite work time, the share of onsite subsidy (tax) in gross income is:

$$SubsidySh_{is}^{exttax} = \frac{t_{is}\theta_{is}^{*}}{1 + t_{is}\theta_{is}^{*}}, \tag{G.2}$$

$$\text{where } t_{is} = \frac{1}{2\theta_{is}^{*2}}\left[-1 + \theta_{is}^{*}(1-\theta_{is}^{*})G_{is}^{*} + \sqrt{(1-\theta_{is}^{*}(1-\theta_{is}^{*})G_{is}^{*})^{2} + 4\theta_{is}^{*2}G_{is}^{*}(1-\theta_{is}^{*})}\right], \tag{G.3}$$

$$G_{is}^{*} = \frac{\widetilde{Z}_{is}e^{kt_i} - \zeta_{is}}{d_{is}(\theta_{is}^{*})\theta_{is}^{*} + \widetilde{Z}_{is}d_{is}(\theta_{is}^{*})(1-\theta_{is}^{*})}. \tag{G.4}$$

In equation (G.1), (G.2), employment and share of time spent working onsite $L_{is}^{m*}, \theta_{is}^{m*}$ refers to the socially optimal labor allocation.

H Data

This section shows the details of deriving residual wages by city-sector-work mode, summary statistics of data, and recovered shifters.

Residual Wages

To obtain the residual wages, I firstly run the mincer regression to obtain individual residual wages for each period:

$$\ln(\mathrm{W}_{\omega}) = c + \boldsymbol{\beta X_{\omega}} + \epsilon_{\omega}, \tag{H.1}$$

where $ln(\mathrm{W}_{\omega})$ is the log of individual-level hourly wages (2020 dollars). $\boldsymbol{X_{\omega}}$ are control variables including 3 groups: (1) demographic variables: age, age squared, education, education squared, experience, experience squared, race, gender, marriage status; (2) variables used to control the sorting effect of WFH: number of own children under age 5, Dingel and Neiman (2020)'s WFH feasibility index at detailed occupation level, interaction terms of this index and education, experience, age, and marriage status, respectively; (3) other variables that affects the wages: industry, occupation, year fixed effect. Next, I calculate the average residual wages by averaging residual wages by city, sector, work mode, and months:

$$avgResigualWage_{ism} = \frac{1}{T} \sum_t \left(\frac{1}{L_{ism}} \sum_{\omega \in \{i,s,m\}} \hat{\epsilon}_{\omega(ismt)} \right), \tag{H.2}$$

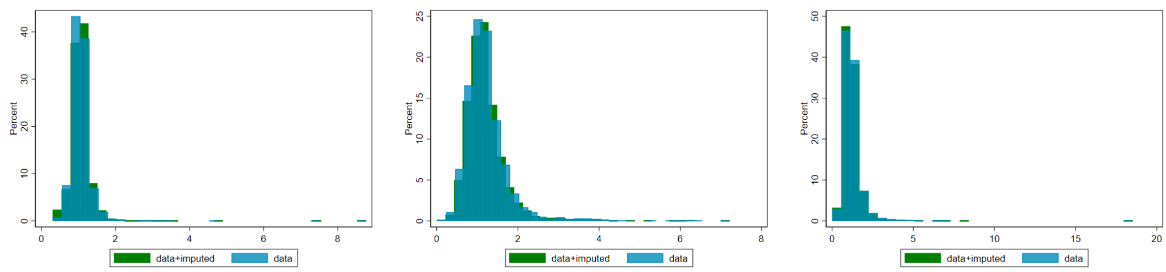
where L_{ism} is the employment at a city-sector-work mode cell, T is the total number of months.

Finally, for imputing wages for city-sector-work mode cells that have positive employment, I run the following regressions:

$$avgResidualWage_{ism} = c + d_{is} + d_{im} + d_{sm} + \epsilon_{ism} \tag{H.3}$$

$$avgResidualWage_{ism} = c + d_i + d_s + d_m + \epsilon_{ism}, \tag{H.4}$$

where $residualWage_{ism}$ is the average residual wages at city-sector-work mode level. c is a constant. d_{is} refers to the 2-way fixed effect at the city-sector level. d_i refers to the city-level fixed effect. The imputed wages use fitted values of average residual wages ($avg\widehat{ResidualW}age_{ism} = \hat{c} + \hat{d}_{is} + \hat{d}_{im} + \hat{d}_{sm}$) from the regression (H.3) if the fitted values is positive. If the fitted values from the regression (H.3) are negative, then imputed wages using the fitted value of residual wages from regression (H.4). Figure 18 compares the wage distribution before and after imputation.



(a) onsite (b) hybrid (c) fully WFH

Figure 18: Distribution of Data Wages and Imputed Wages

Note: If a city-sector cell has positive employment but missing wages, I impute wages using the fitted values of fixed effect regressions. If there is zero employment in a city-sector-work mode cell, the corresponding wage is set as nan.

Summary Statistics of Data

Table 7: Summary Statistics of Variables Used for Recovering Shifters

	Observations	Mean	S.D.	Min	Median	Max
Onsite Wages	3232	1.06	0.31	0.31	1.04	8.77
Hybrid Wages	2512	1.21	0.55	2.77e-03	1.12	7.22
Fully WFH Wages	2415	1.22	0.66	5.96e-03	1.12	18.52
Onsite Emloyment	3249	2.95e+04	7.91e+04	0	8612.39	2.29e+06
Hybrid Emloyment	3249	5465.46	1.35e+04	0	2845.34	2.80e+05
Fully WFH Emloyment	3249	5308.78	1.24e+04	0	2731.42	1.84e+05
Hybrid Workers' Share of Time Working Onsite	2512	0.60	0.15	0.08	0.59	0.98
Commuting time	260	21.25	3.32	13.29	20.84	35.72

I Numerical Algorithms

I.1 Recovering endogenous productivities and shifters

This algorithm recovers endogenous productivities and other shifters without making assumptions about the function form of the externalities or knowing the parameter of the productivity externalities (λ, λ^R, τ). Specifically, given parameters (σ, η, ρ, k) and variables in the data (employment $L^o_{is}, L^h_{is}, L^f_{is}$, wage $W^o_{is}, W^h_{is}, W^f_{is}$, share of time hybrid workers spend working onsite θ_{is}, commuting time t_i), I recover the endogenous productivity (A_{is}, B_{is}) and the structural shifters $(\bar{y}_{is}, \beta_{is}, \epsilon^o_{is}, \epsilon^h_{is}, \epsilon^f_{is}, \zeta_{is})$ as follows:

1. Guess $A_{is}, B_{is}, \widetilde{\Phi}_i, \beta_{is}$.

2. Calculate total output by city using wages and employment in all sectors and work modes:
$$Y_i = \frac{1}{P} \sum_s \sum_m W^m_{is} L^m_{is}, \tag{I.1}$$
where the price index is normalized $P = 1$.

3. Calculate the city-sector-specific productivity shifter:
$$\bar{y}_{is} = \frac{W_{is}}{P} \left(\frac{y_{is}}{Y_i} \right)^{\frac{1}{\eta}} \frac{1}{y_{is}}, \tag{I.2}$$
where city-sector specific total wages $W_{is} = \sum_m W^m_{is} L^m_{is}$, $y_{is} = B_{is} L^o_{is} + \beta_{is} \ell_{is}(\theta_{is}) L^h_{is} + A_{is} L^f_{is}$, $\ell_{is}(\theta_{is}) = [(A_{is}(1-\theta_{is}))^{\frac{\rho-1}{\rho}} + (B_{is}\theta_{is})^{\frac{\rho-1}{\rho}}]^{\frac{\rho}{\rho-1}}$.

4. Calculate the city-sector-specific unit wages implied in the data:

$$w_{is} = P\bar{y}_{is}\left(\frac{Y_i}{y_{is}}\right)^{\frac{1}{\eta}} = \frac{W_{is}}{y_{is}}. \tag{I.3}$$

5. Recover exogenous hybrid productivity using hybrid wages:

$$\beta_{is} = \frac{W^h_{is}}{\ell_{is}(\theta_{is})w_{is}}. \tag{I.4}$$

6. Recover the onsite productivity using onsite wage:

$$B_{is,new} = \frac{W^o_{is}}{w_{is}}. \tag{I.5}$$

If a city-sector cell does not have onsite workers but has hybrid workers, normalize $\beta_{is} = 1$ and recover onsite productivity according to hybrid worker's wages ($\ell_{is}(\theta_{is}) = \frac{W^h_{is}}{w_{is}}$) and the definition of $\ell_{is}(\theta_{is})$:

$$B_{is,new} = \left[\left(\frac{W^h_{is}}{w_{is}}\right)^{\frac{\rho-1}{\rho}} - (A_{is}(1-\theta_{is}))^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}} \frac{1}{\theta_{is}}. \tag{I.6}$$

Recovering B_{is} using this euqation requires $\frac{W^h_{is}}{w_{is}} > A_{is}(1-\theta_{is})$. If the previous two methods fail to recover onsite productivity and the city-sector cell has hybrid workers, update onsite productivity as follows[22]:

$$B_{is,new} = \bar{B}_i + \bar{B}_s = \frac{1}{S}\sum_s \widetilde{B}_{is} + \frac{1}{N}\sum_i \widetilde{B}_{is}, \tag{I.7}$$

where $\widetilde{B}_{is}$ is $B_{is,new}$ recovered from previous 2 methods.

7. Recover remote productivity using fully WFH wages:

$$A_{is,new} = \frac{W^f_{is}}{w_{is}}. \tag{I.8}$$

[22] The idea is using the fitted value of a fixed effect regression to recover the missing B_{is}. Consider the regression $B_{is} = d_i + d_s + e_{is}$, where d_i, d_s are city and sector fixed effects, respectively. e_{is} is the residual. The fitted value of the regression is $\hat{B}_{is} = \hat{d}_i + \hat{d}_s = \frac{1}{S}\sum_s B_{is} + \frac{1}{N}\sum_i B_{is}$.

Similarly, if a city-sector cell does not have fully WFH workers but has hybrid workers, normalize $\beta_{is} = 1$ and recover remote productivity using hybrid workers' wages and the definition of $\ell_{is}(\theta_{is})$

$$A_{is,new} = \left[\left(\frac{W_{is}^h}{w_{is}} \right)^{\frac{\rho-1}{\rho}} - (B_{is}\theta_{is})^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} \frac{1}{1-\theta_{is}}. \tag{I.9}$$

Recovering A_{is} using this euqation requires $\frac{W_{is}^h}{w_{is}} > B_{is}\theta_{is}$. If A_{is} is not available from previous methods and a city-sector cell has hybrid and fully WFH workers, recover A_{is} using the average index:

$$A_{is,new} = \bar{A}_i + \bar{A}_s = \frac{1}{S} \sum_s \widetilde{A}_{is} + \frac{1}{N} \sum_i \widetilde{A}_{is}. \tag{I.10}$$

where $\widetilde{A}_{is}$ is $A_{is,new}$ recovered from previous 2 methods.

8. The first-order condition of the indirect utility function with respect to θ_{is} is $\frac{1-\theta_{is}}{\theta_{is}} = \left(\frac{A_{is}}{B_{is}}\right)^{\rho-1} (\frac{e^{kt_i}}{\zeta_{is}})^{\rho}$. Given B_{is} and A_{is}, recover WFH amenity cost ζ_{is} based on this equation:

$$\zeta_{is} = \left[\frac{\left(\frac{A_{is}}{B_{is}}\right)^{\rho-1} (e^{kt_i})^{\rho}}{\frac{1-\theta_{is}}{\theta_{is}}} \right]^{\frac{1}{\rho}}. \tag{I.11}$$

In the following cases, the WFH amenity cost cannot be recovered using equation (I.11): (1) a city-sector does not have hybrid workers so θ_{is} in the city-sector is not available; (2) a city-sector does not have fully WFH workers and the wage ratio between hybrid and onsite workers do not satisfy $\frac{W_{is}^h}{W_{is}^o} > \theta_{is}$, therefore A_{is} is not available. In these cases, use the following index ($\tilde{\zeta}_{is}$) as the WFH amenity cost:

$$\tilde{\zeta}_{is} = \bar{z}_i + \bar{z}_s = \frac{1}{S} \sum_s \zeta_{is} + \frac{1}{N} \sum_i \zeta_{is}. \tag{I.12}$$

9. Recover labor supply shifter based on the labor supply equation (29):

$$\epsilon_{is}^m = \frac{d_{is}^m(\theta_{is}^m)}{W_{is}^m} \left(\frac{L_{is}^m}{\bar{L}_i} \right)^{\frac{1}{\sigma}} \widetilde{\Phi}_i, \tag{I.13}$$

where $d^m_{is}(\theta^m_{is}) = e^{kt_i}\theta^m_{is} + \zeta_{is}(1-\theta^m_{is}).\theta^o_{is} = 1, \theta^h_{is} = \theta_{is} \in (0,1)$, and $\theta^f_{is} = 0$.

10. Update $\widetilde{\Phi}_i$:

$$\widetilde{\Phi}_{i,new} = \frac{\Phi_i}{\Phi_{i_1}}, \tag{I.14}$$

where $\Phi_i = \left[\sum_s \sum_m \left(\frac{W^m_{is}\epsilon^m_{is}}{d^m_{is}(\theta^m_{is})}\right)^{\sigma}\right]^{\frac{1}{\sigma}}$.

11. If the error $max\left\{\left|\frac{A_{is,new}}{A_{is}} - 1\right|, \left|\frac{B_{is,new}}{B_{is}} - 1\right|, \left|\frac{\widetilde{\Phi}_{i,new}}{\widetilde{\Phi}_i} - 1\right| \forall i, s\right\}$ is larger than a sufficiently small number, set $x = 0.5x + 0.5x_{new}, x \in \{A_{is}, B_{is}, \widetilde{\Phi}_i\}$, and start from step 2, otherwise stop iteration and obtain the solution for $A_{is}, B_{is}, \widetilde{\Phi}_i$ and the shifters $\bar{y}_{is}, \zeta_{is}, \epsilon^o_{is}, \epsilon^h_{is}, \epsilon^f_{is}$.

Given the value of the externality parameters $(\tau, \lambda, \lambda^R)$, I can further recover the exogenous onsite $\bar{b}_{is}$ and exogenous remote productivity $\bar{a}_{is}$:

$$\bar{b}_{is} = B_{is}\left[L^B_{is}(\theta_{is}) + \tau L^A_{is}(\theta_{is})\right]^{-\lambda}, \quad \bar{a}_{is} = A_{is}[L^A_{is}(\theta_{is}) + \tau L^B_{is}(\theta_{is})]^{-\lambda^R}, \tag{I.15}$$

where $L^B_{is}(\theta_{is}) = L^o_{is} + \theta_{is}L^h_{is}, L^A_{is}(\theta_{is}) = L^f_{is} + (1-\theta_{is})L^h_{is}$.

I.2 The Algorithm for Solving Market Equilibrium

Given parameters $(\sigma, \eta, \rho, k, \tau, \lambda, \lambda^R)$, exogenous variables $(\bar{L}_i, t_i)$, and shifters $(\bar{y}_{is}, \bar{b}_{is}, \bar{a}_{is}, \beta_{is}, \zeta_{is}, \epsilon^o_{is}, \epsilon^h_{is}, \epsilon^f_{is})$, I solve for wages $(W^o_{is}, W^h_{is}, W^f_{is})$, employment $(L^o_{is}, L^h_{is}, L^f_{is})$, and share of time hybrid workers spend working onsite (θ_{is}) using the following algorithm:

1. Guess L^o_{is}, L^h_{is}, and L^f_{is} such that $\sum_s(L^o_{is} + L^h_{is} + L^f_{is}) = \bar{L}_i$. Guess the total wage for workers in different work modes $W^o_{is}, W^h_{is}, W^f_{is}$, share of time hybrid workers spend working onsite θ_{is}, and welfare index $\widetilde{\Phi}_i$.

2. Solve for the share of time spent working onsite for hybrid workers:

$$\theta_{is,new} = \frac{1}{1 + (\frac{e^{kt_i}}{\zeta_{is}})^{\rho}(\frac{A_{is}}{B_{is}})^{\rho-1}}, \tag{I.16}$$

where $A_{is} = \bar{a}_{is}[L^A_{is}(\theta_{is}) + \tau L^B_{is}(\theta_{is})]^{\lambda^R}$, $B_{is} = \bar{b}_{is}\left[L^B_{is}(\theta_{is}) + \tau L^A_{is}(\theta_{is})\right]^{\lambda}$, $L^B_{is}(\theta_{is}) = L^o_{is} + \theta_{is}L^h_{is}$, $L^A_{is}(\theta_{is}) = L^f_{is} + (1-\theta_{is})L^h_{is}$.

3. Solve for the city-sector specific unit wages:

$$w_{is} = P\bar{y}_{is}\left(\frac{Y_i}{y_{is}}\right)^{\frac{1}{\eta}}, \tag{I.17}$$

where $y_{is} = B_{is}L^o_{is} + \beta_{is}\ell_{is}(\theta_{is})L^h_{is} + A_{is}L^f_{is}$. $\ell_{is}(\theta_{is}) = [(A_{is}(1-\theta_{is}))^{\frac{\rho-1}{\rho}} + (B_{is}\theta_{is})^{\frac{\rho-1}{\rho}}]^{\frac{\rho}{\rho-1}}$. $Y_i = \left[\sum_s \left(\bar{y}_{is}(y_{is})^{\frac{\eta-1}{\eta}}\right)\right]^{\frac{\eta}{\eta-1}}$. Update total wages for workers with different work modes:

$$W^o_{is,new} = w_{is}B_{is}, \quad W^h_{is,new} = w_{is}\beta_{is}\ell_{is}(\theta_{is}), \quad W^f_{is,new} = w_{is}A_{is}. \tag{I.18}$$

4. Update employment:

$$L^m_{is,new} = \left(\frac{\phi^m_{is}}{\Phi_i}\right)^{\sigma}\bar{L}_i \tag{I.19}$$

where $\phi^m_{is} = \frac{W^m_{is}\epsilon^m_{is}}{d^m_{is}(\theta^m_{is})}$. $d^m_{is}(\theta^m_{is}) = e^{kt_i}\theta^m_{is} + \zeta_{is}(1-\theta^m_{is})$. $\theta^o_{is} = 1, \theta^h_{is} = \theta_{is} \in (0,1)$, and $\theta^f_{is} = 0$. $\Phi_i = [\sum_s \sum_m (\phi^m_{is})^{\sigma}]^{\frac{1}{\sigma}}$.

5. Define reference city as i_1. Update $\widetilde{\Phi}_i$:

$$\widetilde{\Phi}_{i,new} = \frac{\Phi_i}{\Phi_{i_1}}, \tag{I.20}$$

where $\Phi_i = [\sum_s \sum_m (\phi^m_{is})^{\sigma}]^{\frac{1}{\sigma}}$, $\phi^m_{is} = \frac{W^m_{is}\epsilon^m_{is}}{d^m_{is}(\theta^m_{is})}, m \in \{o,h,f\}$.

6. Check the labor market clearing. Define the error $e^m_{is} = W^m_{is} - W^m_{is,supply}, m \in \{o,h,f\}$, where $W^m_{is,supply}$ is total wages calculated by rearranging the labor supply equation:

$$W^m_{is,supply} = \left(\frac{L^m_{is}}{\bar{L}_i}\right)^{\frac{1}{\sigma}}\frac{\Phi_i d^m_{is}(\theta^m_{is})}{\epsilon^m_{is}} \tag{I.21}$$

7. If $max\left\{\left|\frac{\theta_{is,new}}{\theta_{is}} - 1\right|, \left|\frac{W^m_{is,new}}{W^m_{is}} - 1\right|, \left|\frac{L^m_{is,new}}{L^m_{is}} - 1\right|, \left|\frac{\widetilde{\Phi}_{i,new}}{\widetilde{\Phi}_i} - 1\right|, |e^m_{is}|, \forall i,s,m \in \{h,b\}\right\}$ is larger than a sufficient small number, set set $x = 0.5x + 0.5x_{new}, x \in \{\theta_{is}, W^m_{is}, L^m_{is}, \widetilde{\Phi}_i, e^m_{is}\}$ and start from step 2, otherwise stop iteration and obtain the solution for $\theta_{is}, W^m_{is}, L^m_{is}$.

I.3 The Counterfactual Employment

The following process shows the construction of counterfactual employment, which is the model predicted employment with the onsite and remote productivity for each city fixed at the average levels.

1. Use data and the algorithm in Appendix I.1 to recover the endogenous onsite (B_{is}), remote productivity (A_{is}), as well as other shifters.

2. Calculate the average productivities across cities $\bar{A}_i = \frac{1}{S}\sum_s^S A_{is}$, $\bar{B}_i = \frac{1}{S}\sum_s^S B_{is}$.

3. Give $\bar{A}_i$, $\bar{B}_i$, and other shifters and parameters, solve for counterfactual employment $L_{is}^o, L_{is}^h, L_{is}^f$ using algorithm in I.2. When applying the algorithm, replace A_{is} and B_{is} with fixed values $\bar{A}_i$ and $\bar{B}_i$ rather than solving A_{is} and B_{is} as a function of employment and the share of time spent working onsite.

J Parameters and Shifters for Quantification Analyse

Table 8: Errors for Moment Conditions

	Error
$cov(\varepsilon_{is}^o, \bar{y}_{is}\bar{b}_{is}) = 0$	1.51e-03
$cov(\varepsilon_{is}^f, \bar{y}_{is}\bar{a}_{is}) = 0$	4.16e-04
$cov(\varepsilon_{is}, \beta_{is}\frac{\bar{a}_{is}}{\bar{b}_{is}}) = 0$	-2.30e-03
$cov(\frac{\zeta_{is}}{e^{kt_i}}, \frac{\bar{b}_{is}}{\bar{a}_{is}}) = 0$	8.96e-04
$cov(t_i, \frac{\zeta_{is}\bar{a}_{is}}{\bar{b}_{is}}) = 0$	-7.62e-05

Note: All covariances in this table are weighted by the city-sector cells' employment.

Table 9: Parameters for Quantification Analyse

Parameter	Description	Value	Target/Literature
λ	Extensive margin of onsite network externality	0.088	$cov(\varepsilon_{is}^o, \bar{y}_{is}b_{is}) = 0$
λ^T	Extensive margin of remote network externality	0.068	$cov(\varepsilon_{is}^f, \bar{y}_{is}\bar{a}_{is}) = 0$ and $cov(\varepsilon_{is}, \beta_{is}\frac{\bar{a}_{is}}{\bar{b}_{is}}) = 0$
τ	Cross-mode productivity contribution	0.009	$cov(\frac{\zeta_{is}}{e^{kt_i}}, \frac{\bar{b}_{is}}{\bar{a}_{is}}) = 0$
ρ	Elasticity of Substitution between WFH and onsite work	1.304	$cov(t_i, \frac{\zeta_{is}\bar{a}_{is}}{\bar{b}_{is}}) = 0$
σ	Labor supply elasticity	1.26	Burstein et al. (2019)
η	Labor demand elasticity	0.8	Lichter et al. (2015),Beaudry et al. (2018)
k	Elasticity of commuting costs	0.01	Ahlfeldt et al. (2015)
$1-\alpha$	Share of consumption in housing	0.24	Davis and Ortalo-Magné (2011)
$\frac{\gamma}{1-\gamma}$	price elasticity of housing supplied	1.75	Saiz (2010)

Table 10: Recovered Shifters

	Observations	Mean	S.D.	Min	Median	Max
city-sector sepcifc productivity ($\bar{y}_{is}$)	3249	0.05	0.05	9.95e-05	0.04	0.34
exogenous onsite productivity ($\bar{b}_{is}$)	3239	0.92	0.60	0.11	0.73	11.44
exogenous hybrid productivity (β_{is})	2512	0.13	0.08	8.34e-04	0.12	1.02
exogenous remote productivity ($\bar{a}_{is}$)	2751	1.23	0.75	6.12e-03	1.04	9.02
endogenous total onsite productivity (B_{is})	3239	2.08	1.23	0.19	1.75	22.20
endogenous total remote productivity (A_{is})	2751	2.21	1.35	0.01	1.90	15.90
onsite labor supply shifter (ε^o_{is})	3249	0.12	0.08	0	0.10	0.64
hybrid labor supply shifter (ε^h_{is})	3249	0.04	0.24	0	0.03	13.81
fully WFH labor supply shifter (ε^f_{is})	3249	0.07	0.19	0	0.04	8.33
WFH amenity cost (ζ_{is})	2751	2.36	2.23	0.19	1.73	24.58

K Estimating the Relative Strength of Productivity Spillovers

This section uses the following regressions to estimate the difference between remote and onsite spillover elasticity.

$$\ln(\mathrm{W}_\omega) = c + \boldsymbol{\beta X_\omega} + \epsilon_\omega, \tag{K.1}$$

$$\ln\left(\frac{\mathrm{W}^{WFH}_{is}}{\mathrm{W}^{onsite}_{is}}\right) = (\lambda^R - \lambda) ln(L_{is}) + \beta_i + \beta_s + \epsilon_{is}, W^m_{is} = \frac{1}{T}\frac{1}{N^m_{ist}} \sum_{\omega \in \{i,s,t\}} \hat{\epsilon}^m_\omega, \tag{K.2}$$

where equation (K.1) produces the individual residual wages ϵ_ω. W^m_{is} is the average residual wages for onsite or remote workers in city i and industry s. The set of control variables $\boldsymbol{X_\omega}$ consists of 3 groups: (1) demographic variables: age, age squared, education, education squared, experience, experience squared, race, gender, marriage status; (2) variables used to control the sorting effect of WFH: number of own children under age 5, Dingel and Neiman (2020)'s WFH feasibility index at detailed occupation level (aggregated level if the detailed level index is not available), interaction terms of this index and education, experience, age, and marriage status, respectively; (3) other variables that affects the wages: industry, occupation, year fixed effect.

The instrumental variables used to identify $\lambda^R - \lambda$ are the sum of labor supply shifters or counterfactual employment by work mode for each city and sector. Labor supply shifters are recovered using the algorithm in I.1. Counterfactual employment is obtained using the algorithm in I.3.

Using counterfactual employment as a model-implied IV is similar to the approach in Rossi-Hansberg et al. (2019). In the model, employment and wages are simultaneously decided. The increase in employment in a sector or work mode leads to an increase in productivity and wages, which attracts more workers to choose the sector or work mode. Counterfactual employment is the model-predicted employment with the onsite and remote productivity for each city fixed at the average levels. By fixing productivity, I control the channel that higher wages lead to higher employment in a sector or work mode.

The following regression results indicate that most linear estimates of the difference between remote and onsite agglomeration elasticities range between -0.06 and 0.

Table 11: Relative Strength of Remote and Onsite Spillovers

Panel A: CPS data

	OLS	IV: labor supply shifter				IV: counterfactural employment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	ln(WFH wage premium)	ln(employment)	ln(WFH wage premium)	ln(employment)	ln(WFH wage premium)	ln(employment)	ln(WFH wage premium)	ln(employment)	ln(WFH wage premium)
ln(employment)	0.0403***		-0.0336**		-0.2467***		0.0063		-0.0367
	(0.0061)		(0.0158)		(0.0576)		(0.0126)		(0.0267)
ln(labor supply shifter)		1.2085***		0.8263***					
		(0.0181)		(0.0462)					
ln(counterfactal employment)						1.0155***		0.9742***	
						(0.0024)		(0.0047)	
City FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	No	No	Yes	Yes	No	No	Yes	Yes
Obs.	2,401	2,401	2,401	2,401	2,401	2,401	2,401	2,401	2,401
R^2	.032								
K.P.F.		4,466	.	320	.	181,185	.	42,929	.

Panel B: ACS data

	OLS	IV: labor supply shifter				IV: counterfactural employment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	ln(WFH wage premium)	ln(employment)	ln(WFH wage premium)	ln(employment)	ln(WFH wage premium)	ln(employment)	ln(WFH wage premium)	ln(employment)	ln(WFH wage premium)
ln(employment)	0.0126***		-0.0079		-0.0557***		-0.0074		-0.0420***
	(0.0032)		(0.0077)		(0.0188)		(0.0073)		(0.0148)
ln(labor supply shifter)		1.2270***		0.7218***					
		(0.0150)		(0.0406)					
ln(counterfactal employment)						1.0417***		0.8294***	
						(0.0058)		(0.0166)	
City FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	No	No	Yes	Yes	No	No	Yes	Yes
Obs.	7,978	3,159	3,159	3,159	3,159	3,159	3,159	3,159	3,159
R^2	.009								
K.P.F.		6,704	.	316	.	32,315	.	2,487	.

In panel A, the data source is the CPS from 2022 to 2024. The WFH wage premium refers to the ratio of residual wages for fully WFH workers to residual wages for fully onsite workers. In panel B, the data source is the ACS from 2021 to 2023. The WFH wage premium refers to the ratio of residual wages for home-based workers to residual wages for commuting workers. Robust standard errors are shown in parentheses. $p < 0.1, **p < 0.05, ***p < 0.01$. Regressions are weighted by employment size at the CBSA-industry level.

References

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The Economics of Density: Evidence From the Berlin Wall. *Econometrica 83*(6), 2127–2189.

Barrero, J., N. Bloom, and S. Davis (2023). The Evolution of Work from Home. *Journal of Economic Perspectives 37*(4), 23–50.

Barrero, J. M., N. Bloom, and S. J. Davis (2021, may). Why Working From Home Will Stick. *National Bureau of Economic Research Working Paper 28731*.

Beaudry, P., D. A. Green, and B. M. Sand (2018). In Search of Labor Demand. *American Economic Review 108*(9), 2714–2757.

Behrens, K., S. Kichko, and J.-F. Thisse (2021). Working From Home: Too Much of a Good Thing? *CESifo Working Paper*.

Bertram, J., C. S. Ruhnke, and J. R. Sch¨ondube (2024). Optimal Degree of Remote Work.

Brucks, M. S. and J. Levav (2022). Virtual communication curbs creative idea generation. *Nature 605*(7908), 108–112.

Burstein, A., E. Morales, and J. Vogel (2019). Changes in Between-Group Inequality: Computers, Occupations, and International Trade. *American Economic Journal: Macroeconomics 11*(2), 348–400.

Combes, P. P. and L. Gobillon (2015). The Empirics of Agglomeration Economies. In *Handbook of Regional and Urban Economics*, Volume 5, pp. 247–348.

Davis, M. A., A. C. Ghent, and J. Gregory (2024). The Work-from-Home Technology Boon and Its Consequences. *Review of Economic Studies*, rdad114.

Davis, M. A. and F. Ortalo-Magné (2011, apr). Household expenditures, wages, rents. *Review of Economic Dynamics 14*(2), 248–261.

Delventhal, M. and A. Parkhomenko (2023). Spatial Implications of Telecommuting. *Available at SSRN: https://ssrn.com/abstract=3746555 or http://dx.doi.org/10.2139/ssrn.3746555*.

Delventhal, M. J., E. Kwon, and A. Parkhomenko (2022). How do cities change when we work from home? *Journal of Urban Economics 127*.

Ding, Y. and M. Ma (2023). Return-to-office mandates. *URL: https://papers. ssrn. com/abstract 4675401.*

Dingel, J. and B. Neiman (2020, apr). How Many Jobs Can be Done at Home? *Journal of Public Economics 189*, 104235.

Emanuel, N., A. Pallais, and E. Harrington (2022). The Power of Proximity.

Fajgelbaum, P. D. and C. Gaubert (2020). Optimal Spatial Policies, Geography, and Sorting. *The Quarterly Journal of Economics 135*(2), 959–1036.

Farrokhi, F. (2021). skill, agglomeration, and inequality in the spatial economy. *International Economic Review 62*(2), 671–721.

Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, D. Backman, A. Chen, G. Cooper, S. Richards, M. Schouweiler, and M. Westberry (2024). IPUMS CPS: Version 12.0 [dataset]. Accessed: 2025-02-06.

Flynn, S., A. C. Ghent, and V. Nair (2024). Determinants and Consequences of Return to Office Policies. *SSRN Electronic Journal*.

Forman, C. and N. Van Zeebroeck (2012). From wires to partners: How the internet has fostered R & D collaborations within firms. *Management Science 58*(8), 1549–1568.

Forman, C. and N. van Zeebroeck (2019, oct). Digital technology adoption and knowledge flows within firms: Can the Internet overcome geographic and technological distance? *Research Policy 48*(8), 103697.

Frakes, M. D. and M. F. Wasserman (2021, jun). Knowledge Spillovers, Peer Effects, and Telecommuting: Evidence from the U.S. Patent Office. *Journal of Public Economics 198*, 104425.

Harrington, E. and N. Emanuel (2021). Working' Remotely? Selection, Treatment, and Market Provision of Remote Work (JMP). *Working paper*.

Heblich, S., S. J. Redding, and D. M. Sturm (2020, nov). The making of the modern metropolis: Evidence from London. *Quarterly Journal of Economics 135*(4), 2059–2133.

Lennox, J. (2020). More working from home will change the shape and size of cities. *Centre of Policy Studies/IMPACT Centre Working Papers*.

Lichter, A., A. Peichl, and S. Siegloch (2015). The Own-Wage Elasticity of Labor Demand: A Meta-Regression Analysis. *European Economic Review 80*, 94–119.

Liu, S. and Y. Su (2022). The Effect of Working from Home on the Agglomeration Economies of Cities: Evidence from Advertised Wages. *SSRN Electronic Journal*.

Melo, P. C., D. J. Graham, and R. B. Noland (2009, may). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics 39*(3), 332–342.

Monte, F., C. Porcher, and E. Rossi-hansberg (2023). Remote Work and City Structure. *NBER Working Paper*.

Pollard, E. (2019). A New Approach to Industry and Occupation Recoding in the CPS. *Federal Reserve Bank of Kansas City Technical Briefing 19-02*.

Presidente, G. and C. B. Frey (2023). *Disrupting Science.*

Redding, S. J. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*, 21–58.

Rhee, H. J. (2008, jan). Home-based telecommuting and commuting behavior. *Journal of Urban Economics 63*(1), 198–216.

Rossi-Hansberg, E., P.-D. Sarte, and F. Schwartzman (2019). Cognitive Hubs and Spatial Redistribution.

Rossi-Hansberg, E., P.-D. Sarte, and F. Schwartzman (2021, may). Local Industrial Policy and Sectoral Hubs. *AEA Papers and Proceedings 111*, 526–531.

Ruggles, S., S. Flood, M. Sobek, D. Backman, A. Chen, G. Cooper, S. Richards, R. Rodgers, and M. Schouweiler (2024). IPUMS USA: Version 15.0 [dataset]. Accessed: 2025-02-06.

Safirova, E. (2002, jul). Telecommuting, traffic congestion, and agglomeration: A general equilibrium model. *Journal of Urban Economics 52*(1), 26–52.

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics 125*(3), 1253–1296.