The Colocation Friction:
Dual-Earner Job Search and Labor Market Outcomes

Hanno Foerster* Robert Ulbricht†

July 9, 2023

Abstract

Dual-earner households face a colocation problem: They need to find two jobs in one location. We develop a spatial directed search model that captures the unique frictions that characterize the job search by dual-earner households. We derive general conditions under which these "colocation frictions" are binding and quantify their consequences for the U.S. labor market. Estimated at the commuting zone level, the model implies that colocation frictions disproportionately affect women, reducing the earning gains from migration by 27%. The colocation friction strongly discourages migration, especially among "power couples", preempting relocation to more productive and higher-amenity locations in the long-run. Taken together, we estimate that the colocation friction incurs a lifetime utility loss equivalent to a 1.61% decrease in lifetime earnings.

Keywords: Coordinated matching, directed search, dual-earner job search, migration, multiple job applications

JEL Classification: E24, E32, J24, J64.

*Boston College, hanno.foerster@bc.edu
†Boston College, ulbricht@bc.edu

1 Introduction

A substantial share of U.S. workers are part of a dual-earner household. Compared to single-earner households, dual-earner households face a unique constraint on their job search and labor mobility: They need to find two jobs in one location. While largely ignored by previous literature, the constraint is all but certain to shape the labor market experience of workers in dual-earner households (Mincer 1978; Guler, Guvenen & Violante 2012).

In this paper, we develop a spatial search model of the labor market for dual-earner households, that we use to formalize, characterize, and quantify the unique frictions faced by dual-earner households in their job search. Building on Menzio & Shi (2010, 2011), search is directed, allowing household members to coordinate their search effort towards the same locations. Yet, due to the usual labor market frictions, not every application generates a match, and matching succeeds independently across spouses. This exposes dual-earner households to a risk of job offers being spatially mismatched across spouses, which we term the "colocation friction", and which reduces the odds of obtaining a joint job offer within the same location despite coordinating search efforts. This colocation friction is the key friction distinguishing the job search of dual-earner households from the one by single earners.

The colocation friction manifests itself in several ways: First, migration is more likely to create a trailing spouse with possibly long-lasting effects on their career. Second, as a consequence, migration may become less attractive, reducing labor mobility. Third, the colocation friction may affect to which locations dual-earner households direct their search and, therefore, which locations they typically migrate to. Fourth, by influencing labor mobility, workers' location choice, and their labor market trajectories upon migration, the colocation friction may have long-run effects on workers' productivity, employment and welfare.

To measure the prevalence and consequences of the colocation friction, we develop a fictitious benchmark where spouses are counterfactually allowed to correlate their matching success, effectively shutting down the colocation friction, while otherwise being exposed to the same search frictions as in the baseline economy.[1] Equipped with the benchmark, we first apply it to assess under which conditions the colocation friction is binding. We show that it does if and only if a household's lifetime utility is convex in the number of simultaneous job offers in a given location. As we demonstrate, this is more likely for households with two employed spouses, when job ladders are steep, when migration costs are large, and when the search elasticity across locations is small. Conversely, the colocation friction is likely to be slack for households with two non-employed spouses and, when childcare costs are large, for

[1] While we only utilize the benchmark as a measuring device, it can be implemented as part of a decentralized equilibrium by allowing workers with identical job qualifications to trade job offers. In that sense, one may interpret the colocation friction as a market inefficiency.

households with children.

We quantify the colocation friction in a version of our model where locations correspond to U.S. commuting zones and households are characterized by each spouse's occupation, human capital, and whether there are children in the household. Human capital accumulates through learning-by-doing and is estimated to match the empirical steepness of job ladders as in, e.g., Jung & Kuhn (2019) and Jarosch (2022). Commuting zones (CZ) are distinguished by th. ge earnings by 2% and men's average earnings by 0.8%. In addition to its direct effect on post-migration employment, these long-term effects further reflect human capital losses from a lack in work-experience and forgone migration opportunities to more productive commuting zones.

Finally, taken together, we estimate that the colocation friction incurs a lifetime utility loss equivalent to a decrease in lifetime earnings ranging from 0.5% to 2.7%, depending on both spouses' occupations.

Related literature The contribution of this paper is threefold. First, we develop a novel spatial search model for studying dual-earner households' job search and migration behavior. Our model builds on the directed search models by Menzio & Shi (2010, 2011), Menzio, Telyukova & Visschers (2016), Schaal (2017), and Herkenhoff, Phillips & Cohen-Cole (2022), which we extend to dual-earner search and to add a spatial dimension. For modeling dual-earner search, adopting a directed search rather than a random search approach is crucial as it reflects that spouses may coordinate their job search towards the same locations. Moreover, we benefit computationally from block recursivity (Menzio & Shi 2010; Wright, Kircher, Julien & Guerrieri 2021), which makes our quantitative model tractable despite a large state space and action space.

Second, we contribute to the literature on labor misallocation. Previous studies have explored various sources of misallocation on labor markets, including search frictions (Hosios 1990) and monopsony power (Galenianos, Kircher & Virág 2011; Rabinovich & Wolthoff 2022), spatial frictions (Şahin, Song, Topa & Violante 2014; Findeisen, Lee, Porzio & Dauth 2021), and information frictions (Jovanovic 1979, 1984). Our paper contributes a novel aspect to this literature by exploring the unique frictions characterizing dual-earner job search across multiple locations and by quantifying their consequences for the allocation of workers to jobs.

Third, we add to a small literature that has analyzed dual-earner households' labor market and migration decisions.[2] This literature goes back to Mincer (1978), and has explored how

[2]This literature is at the intersection of a somewhat larger literature studying dual-earner job search (e.g, Dey & Flinn 2008; Guler et al. 2012; Pilossoph & Wee 2021; Flabbi & Mabli 2018) and the literature on domestic migration, which typically focuses on single-earners (e.g., Kennan & Walker 2011; Kaplan & Schulhofer-Wohl 2017; Piyapromdee 2020).

migration decisions are impacted by household composition, as well as gender-specific labor market opportunities (Costa & Kahn 2000; Gemici 2007), and how policy affects dual-earner households' migration decisions (Venator 2021). Relative to this literature our paper is the first to model dual-earner job search as directed, which arguably is crucial in our context as it allows spouses to coordinate their search towards specific locations. Unlike random search models, our framework thus accounts for coordination in search efforts, which allows us to formalize and quantify the fundamental friction arising from a spatial mismatch in matching successes. Leveraging our framework, we provide general results on when the risk of spatially mismatched job offers constrains dual-earner job search, and quantify the implications for the U.S. labor market.

Layout The paper proceeds as follows. Section 2 introduces the general framework. Section 3 characterizes the colocation friction, develops our measuring approach, and presents a simple example to illustrate the economics behind the friction. Section 4 introduces and calibrates the quantitative model. Section 5 studies the consequences of the colocation friction for employment, earnings and welfare. Section 6 concludes.

2 General Framework

We develop a spatial search model of the labor market for dual-earner households. Search is directed, allowing household members to coordinate their search effort towards the same locations. Yet, due to labor market frictions, matching success is random, reducing the odds that during any given time frame both spouses successfully match within the same location.

2.1 Environment

Preferences and technology Time is continuous and extends forever. There is a finite set of commuting zones, indexed by $r \in \mathcal{R}$. The economy is populated by an endogenous measure of one-vacancy firms and a unit measure of households. Each household consists of two adult workers or "spouses", indexed by $i \in \{1, 2\}$. While later, in our quantification, i maps into workers' "gender", there is no need for now to make assumptions about the gender-composition within households. Firms and households are risk neutral and share the same effective discount rate ρ.

Following the literature, we assume that search and matching is privately efficient. In order to characterize labor and migration flows, it then suffices to specify the sum of a household's instantaneous utility flow and the labor product of its employed members. Let

$$u(\boldsymbol{e}, \boldsymbol{s}, r) = \bar{u}(\boldsymbol{e}, \boldsymbol{s}, r) + \sum_{i \in \{1,2\}} z_i(\boldsymbol{s}, r) \cdot \mathbb{1}_{e_i=1} \tag{1}$$

denote this joint value flow. Here, $\boldsymbol{e} \equiv (e_1, e_2) \in \{0,1\}^2$ is the employment status of the household's adult members, $\bar{u}$ is their utility flow net of earnings, z_i is the labor product of spouse i, $r \in \mathcal{R}$ is the household's current location, and $\boldsymbol{s} \in \mathcal{S}_1 \times \cdots \times \mathcal{S}_{n_s}$ is a generic "catch-all" state that captures other persistent and transitory characteristics of the household. For example, in our quantification, $\boldsymbol{s}$ includes the occupation of both spouses, their human capital, and whether or not they have children. We assume that $\boldsymbol{s}$ has finite support $\mathcal{S}_k$ in all its dimensions $k \in \{1, \ldots, n_s\}$.

Other than through search and migration, a household's type $(\boldsymbol{e}, \boldsymbol{s}, r)$ evolves stochastically with Poisson arrival rates given by $\lambda(\boldsymbol{e}', \boldsymbol{s}', r' | \boldsymbol{e}, \boldsymbol{s}, r)$. In our quantification, we specify λ to expose households to exogenous job separations, human capital dynamics, location preference shocks, and the arrival and departure of children.

Labor markets and migration The labor market is organized in a continuum of submarkets indexed by the location of jobs $q \in \mathcal{R}$, the worker type $(i, \boldsymbol{s})$, and the firm's share y of the joint value of the match.[3] Workers direct their search toward these submarkets, choosing both a firm share y and a search effort $\kappa_{i,q}$ for each location q. Specifically, each spouse $i \in \{1,2\}$ is endowed with a type-specific search budget $\bar{\kappa}_i(\boldsymbol{e}, \boldsymbol{s})$, which they can allocate to search across submarkets in different locations subject to

$$\left(\sum_{q \in \mathcal{R}} \kappa_{i,q}^{\frac{1+\eta}{\eta}} \right)^{\frac{\eta}{1+\eta}} \leq \bar{\kappa}_i(\boldsymbol{e}, \boldsymbol{s}). \tag{2}$$

Here, $\eta \geq 0$ is the elasticity of substitution between locations. In the limit where $\eta \to \infty$, workers allocate their entire search budget to the single location with the highest gains from search as is usually the case in directed search models. At the other extreme where $\eta = 0$, diversifying search is costless and workers allocate $\bar{\kappa}_i(\boldsymbol{e}, \boldsymbol{s})$ units of search effort to each location with positive search gains as in the literature on multiple job applications (Albrecht, Gautier & Vroman 2006; Kircher 2009; Galenianos & Kircher 2009).

Vacancies are created by an infinite supply of potential firms, which can open vacancies in any submarket at flow costs c. Vacancies and workers in a given submarket come together

[3] Indexing submarkets by firms' share y is equivalent to indexing by workers' lifetime utility along with household types $(\boldsymbol{e}, \boldsymbol{s}, r)$. It is worth noting that while indexing by y yields broader submarkets, it is isomorphic to finer partitions as long as vacancies are created at constant returns to scale (as we indeed impose below).

through a frictional matching process. In particular, a worker searching in submarket $\psi \equiv (q, y, i, \boldsymbol{s})$ meets a vacancy at rate $\kappa_{i,q} p(\theta_\psi)$, where θ_ψ is the ratio between vacancies posted and effective search effort allocated to submarket ψ. Similarly, a vacancy posted in submarket ψ meets a worker at rate $p(\theta_\psi)/\theta_\psi$. As usual, we assume that p is twice differentiable, strictly increasing and concave, with $p(0) = p'(\infty) = 0$ and $p'(0) = \infty$.

When a firm and a worker meet, the firm offers a wage contract with present discounted value equal to the match value net of the firm's share y and hires the worker. Following Menzio & Shi (2010, 2011), we assume that the underlying contract space is complete. In particular, endogenous separations and on-the-job search maximize the joint value of the household and all its current employers.

New jobs entail migration whenever the new job is located in a commuting zone q that differs from a household's current location r. In this case, the spouse without the job offer quits their job and the household moves to q. Migration entails a utility cost $\chi(q|\boldsymbol{s}, r)$, normalized so that $\chi(r|\cdot, r) = 0$.

2.2 Equilibrium Characterization

Vacancy creation By free entry, the value of creating a vacancy must be zero in every submarket. From firms' zero profit condition,

$$c\theta_\psi = p(\theta_\psi) \max\{0, y\}, \tag{3}$$

this pins down the market tightness θ_ψ as a function of y.

Search and separation policies Next, consider the search, migration and separation policies of workers and their employers. Private efficiency implies that the policies maximize the joint value of a household and its current employers, given by

$$\begin{aligned}\rho V(\boldsymbol{e}, \boldsymbol{s}, r) = u(\boldsymbol{e}, \boldsymbol{s}, r) &+ \sum_{\boldsymbol{e}', \boldsymbol{s}', r'} \lambda(\boldsymbol{e}', \boldsymbol{s}', r'|\boldsymbol{e}, \boldsymbol{s}, r)\Big(V(\boldsymbol{e}', \boldsymbol{s}', r') - V(\boldsymbol{e}, \boldsymbol{s}, r)\Big) \\ &+ \max_{\{\kappa_{i,q}, y_{i,q}\}} \sum_{i,q} \kappa_{i,q} p(\theta_{i,q})\Big(V(\boldsymbol{e}^{\text{new},i}, \boldsymbol{s}, q) - y_{i,q} - \chi(q|\boldsymbol{s}, r) - V(\boldsymbol{e}, \boldsymbol{s}, r)\Big) \\ &+ \lim_{\pi \to \infty} \pi \sum_i \max\Big\{0, V(\boldsymbol{e}^{\text{sep},i}, \boldsymbol{s}, q) - V(\boldsymbol{e}, \boldsymbol{s}, q)\Big\}. \end{aligned} \tag{4}$$

The joint flow value is comprised of four terms: (i) the instantaneous value flow u as defined in (1); (ii) the value change induced by exogenous shocks to the household type $(\boldsymbol{e}, \boldsymbol{s}, r)$; (iii) the value change induced by either spouse finding a new job, which is maximized subject to the θ–y frontier posed by (3); and (iv) the option for either spouse to quit their job. Here, $\boldsymbol{e}^{\text{new},i}$ is a household's employment status after spouse i accepts a new job, defined by $e_i^{\text{new},i} = 1$ and $e_{-i}^{\text{new},i} = e_{-i} \cdot \mathbb{1}_{q=r}$ where "$-i$" denotes the spouse without job offer. Similarly, $\boldsymbol{e}^{\text{sep},i}$ is the employment status after spouse i quits their job, defined by $e_i^{\text{sep},i} = 0$ and $e_{-i}^{\text{sep},i} = e_{-i}$.

It remains to characterize the optimal choice of $\{\kappa_{i,q}, y_{i,q}\}$ in households' job search. Consider first the value split between households and firms. From (3), the market tightness is increasing in firms' share y, creating a trade-off for workers to search in submarkets with higher job finding rates versus searching in submarkets with higher household shares. Maximizing the third term in (4) subject to the θ–y frontier defined by (3), the optimal market tightness is given by

$$\theta_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r) = p'^{-1}\left(\frac{c}{V(\boldsymbol{e}^{\text{new},i}, \boldsymbol{s}, q) - \chi(q|\boldsymbol{s}, r) - V(\boldsymbol{e}, \boldsymbol{s}, r)}\right), \tag{5}$$

which in turn pins down the matching rate per unit of search effort, $p(\theta_\psi)$.

Finally, consider the allocation of search effort across commuting zones, $\{\kappa_{i,q}\}$. Let $\Lambda_{i,q}$ denote spouse i's expected gain per unit of search effort allocated to location q,

$$\Lambda_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r) \equiv p(\theta^*_{i,q})\Big(V(\boldsymbol{e}^{\text{new},i}, \boldsymbol{s}, q) - \chi(q|r) - V(\boldsymbol{e}, \boldsymbol{s}, r)\Big) - \theta^*_{i,q}c,$$

with $\theta^*_{i,q} \equiv \theta_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r)$ denoting the optimal market tightness in (5). Maximizing the joint value (4) subject to the search budget (2), then yields

$$\kappa_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r) = \Lambda_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r)^\eta \left(\sum_v \Lambda_{i,v}(\boldsymbol{e}, \boldsymbol{s}, r)^{1+\eta}\right)^{-\frac{\eta}{1+\eta}} \bar{\kappa}_i(\boldsymbol{e}, \boldsymbol{s}). \tag{6}$$

Together with (5) this pins down job finding and migration rates,

$$f_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r) \equiv \kappa_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r) \cdot p(\theta_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r)), \tag{7}$$

completing the characterization of search policies.

Steady state equilibrium In this economy, all policy rules are functions of only the idiosyncratic household type $(\boldsymbol{e}, \boldsymbol{s}, r)$. An equilibrium is a collection of maps from $(\boldsymbol{e}, \boldsymbol{s}, r)$ to search and separation policies satisfying (4)–(6) along with a value split satisfying (3). Throughout we focus on the case where the cross-sectional distribution over $(\boldsymbol{e}, \boldsymbol{s}, r)$ is at steady state.[4]

[4] The cross-sectional distribution is characterized by a standard Kolmogorov forward equation, detailed in Appendix A.2.

3 The Colocation Friction

Conditional on search policies, matching in the baseline economy is statistically *independent* across spouses, reducing the odds of obtaining a joint job offer in any given location below the individual matching rates. While originating from search frictions, which lead to randomness in the matching process even for single earners, the risk of job offers being spatially mismatched across spouses is unique to dual-earner households. We refer to this unique risk as "colocation friction". As hypothesized by Mincer (1978), the friction may harm households in two ways. First, generating trailing spouses, it lowers the employment and earnings of migrating households. Second, as a consequence, it may also deter migration in the first place.

To evaluate this hypothesis, we introduce a fictitious benchmark where spouses are counterfactually allowed to *correlate* their individual matching success. In minimizing spatial mismatch across job offers, it effectively shuts down the colocation friction. Otherwise, the benchmark exposes households to exactly the same search frictions as in the baseline economy. Comparing the job search and migration choices between the benchmark and baseline economy, we then characterize the precise conditions under which the colocation friction is binding and, if it is, what its consequences are for employment and earning trajectories.

3.1 Correlated Matching Benchmark

Our benchmark is a replica of the baseline economy in Section 2 with the twist that households may choose to correlate their matching success within commuting zones.

Formally, suppose two spouses allocate search efforts $\{\kappa_{i,q}\}_{i=1,2}$ towards finding jobs in commuting zone q and submarkets with market tightness $\{\theta_{i,q}\}_{i=1,2}$. Let $f_{i,q} = \kappa_{i,q} p(\theta_{i,q})$ denote their individual job finding rates. The correlated matching benchmark is then characterized by a choice of correlated matching rates $\{\omega_q\}_{q \in \mathcal{R}}$ with

$$0 \leq \omega_q \leq \min_{i \in \{1,2\}} \{f_{i,q}\} \quad \text{for all } q \in \mathcal{R}. \tag{8}$$

Given their choice of $\{\omega_q\}$, both spouses obtain a joint job offer in commuting zone q at rate ω_q, and obtain individual offers at the residual rates $f_{i,q} - \omega_q$. If they obtain a joint job offer, both spouses move to employment, $\boldsymbol{e} = (1, 1)$, independently of the location of their new jobs.

We note that the benchmark does not change the primitives of the search technology. That is, workers face the same constraint (2) on their allocation of search effort $\{\kappa_{i,q}\}$, firms face the same cost of vacancy creation c, and matching is subject to the same frictional matching function $p(\theta_\psi)$. When $\omega_q = 0$ for all q, the benchmark yields search and migration policies that are identical to the baseline economy.

3.2 When Is the Colocation Friction Binding?

We say that for a household of type $(\boldsymbol{e}, \boldsymbol{s}, r)$ the colocation friction is binding while searching for jobs in commuting zone q if and only if they would choose to correlate their matching success if given the choice. That is, provided with optimal correlation policies $\{\omega_q\}$ from the fictitious benchmark, we identify the friction to be binding precisely when $\omega_q(\boldsymbol{e}, \boldsymbol{s}, r) > 0$. The following proposition characterizes when this is the case.

Proposition 1. *Fix a joint value function V. Let*

$$\Delta V_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r) = V(\boldsymbol{e}_i^{\text{new},i}, \boldsymbol{s}, q) - \chi(q|\boldsymbol{s}, r) - V(\boldsymbol{e}, \boldsymbol{s}, r)$$

denote the gains from an individual match by spouse i. Similarly, let

$$\Delta V_q^{\text{corr}}(\boldsymbol{e}, \boldsymbol{s}, r) = V(\boldsymbol{e}^{\text{corr}}, \boldsymbol{s}, q) - \chi(q|\boldsymbol{s}, r) - V(\boldsymbol{e}, \boldsymbol{s}, r)$$

denote the gains from a correlated match, where $\boldsymbol{e}^{\text{corr}} = (1, 1)$. Then correlated matching is beneficial if and only if

$$\Delta V_q^{\text{corr}}(\boldsymbol{e}, \boldsymbol{s}, r) \geq \sum_{i \in \{1,2\}} \Delta V_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r). \tag{9}$$

If condition (9) holds, the correlated matching rate is optimally set to its upper bound, $\omega_q = \min_{i \in \{1,2\}} \{f_{i,q}\}$. Otherwise, it is optimally set to its lower bound, $\omega_q = 0$.

The formal proof is in Appendix A.1. Intuitively, as we illustrate in Figure 1, condition (9) assesses the curvature of $V - \chi$ in the number of simultaneous job offers in a given location q. If $V - \chi$ is convex in the number of simultaneous job offers, then the value gain of a joint offer exceeds the average gains of the individual offers and the colocation friction is binding. Conversely, if $V - \chi$ is concave, the household is better off hedging its bets by generating statistically independent offers and the colocation friction is slack.

It is worth noting that Proposition 1 applies to any value function V. In what follows we distinguish two types of application. First, we study for whom the colocation friction is binding in the baseline economy by examining condition (9) for the baseline value function. Second, we quantify the consequences of the colocation friction for the U.S. labor market by counterfactually relaxing it, which involves computing a counterfactual value function for the benchmark economy.

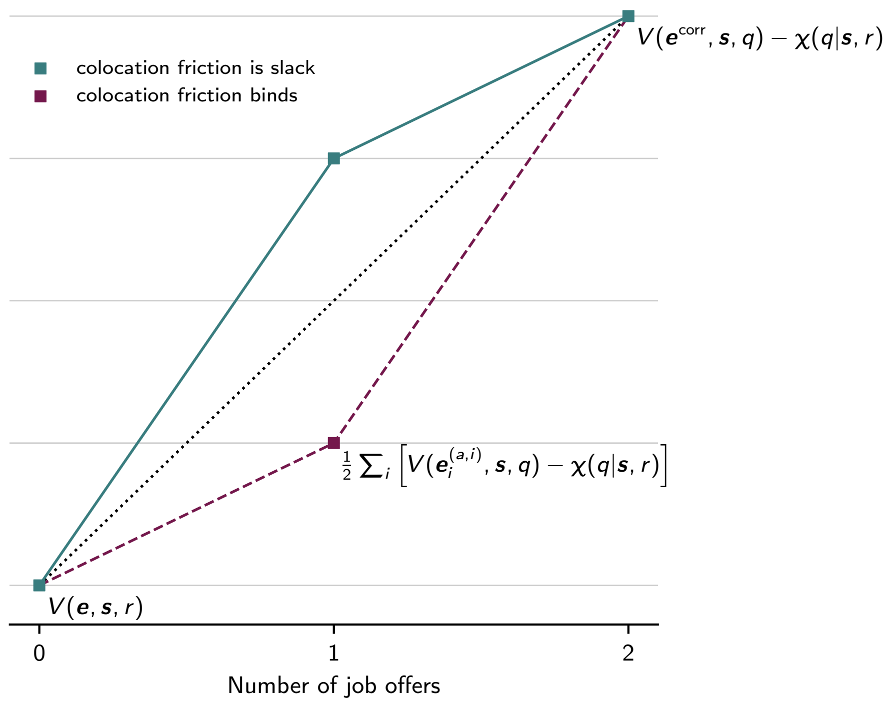

Figure 1: Illustration of condition (9). The colocation friction binds if and only if, averaged across spouses, $V - \chi$ is convex in the number of simultaneous job offers as in the dashed purple graph, and is slack if $V - \chi$ is concave as in the solid turquoise graph.

3.3 Simple Example

Before quantifying our framework, we briefly explore a simple example to illustrate the economic forces that determine the curvature of V. Consider a household with initial type $(\boldsymbol{e}_0, \boldsymbol{s}_0, r_0)$. There are no exogenous shocks, $\lambda(\cdot|\cdot) = 0$, and both spouses and all locations $q \neq r_0$ are symmetric. Specifically, we have $\bar{\kappa}_i(\boldsymbol{e}, \boldsymbol{s}) = \bar{\kappa}$, $\chi(q|\boldsymbol{s}, r) = \chi$, and

$$u(\boldsymbol{e}, \boldsymbol{s}, r) = \begin{cases} u_0 & r = r_0 \\ u_{e_1+e_2} & r \neq r_0, \end{cases}$$

with $u_0 + \rho\chi \leq u_1 \leq u_2$. Finally, we normalize $p(\infty) = 1$ and simplify by considering the limit where the matching function is inelastic in vacancies, $\mathrm{d}\log p/\mathrm{d}\log\theta \to 0$.

The example admits a simple recursive solution where $V(\boldsymbol{e}, \boldsymbol{s}, r) \in \{V_0, V_1, V_2\}$. Specifically, if $r \neq r_0$ and both spouses are employed, we have $\rho V_2 = u_2$. If $r \neq r_0$ and only one spouse is employed, we have $\rho V_1 = u_1 + \bar{\kappa}\,(V_2 - V_1)$. Finally, the initial value at $r = r_0$ is given by

$$\rho V_0 = u_0 + 2\bar{\kappa} N^{\frac{1}{1+\eta}}\,(V_1 - \chi - V_0)\,,$$

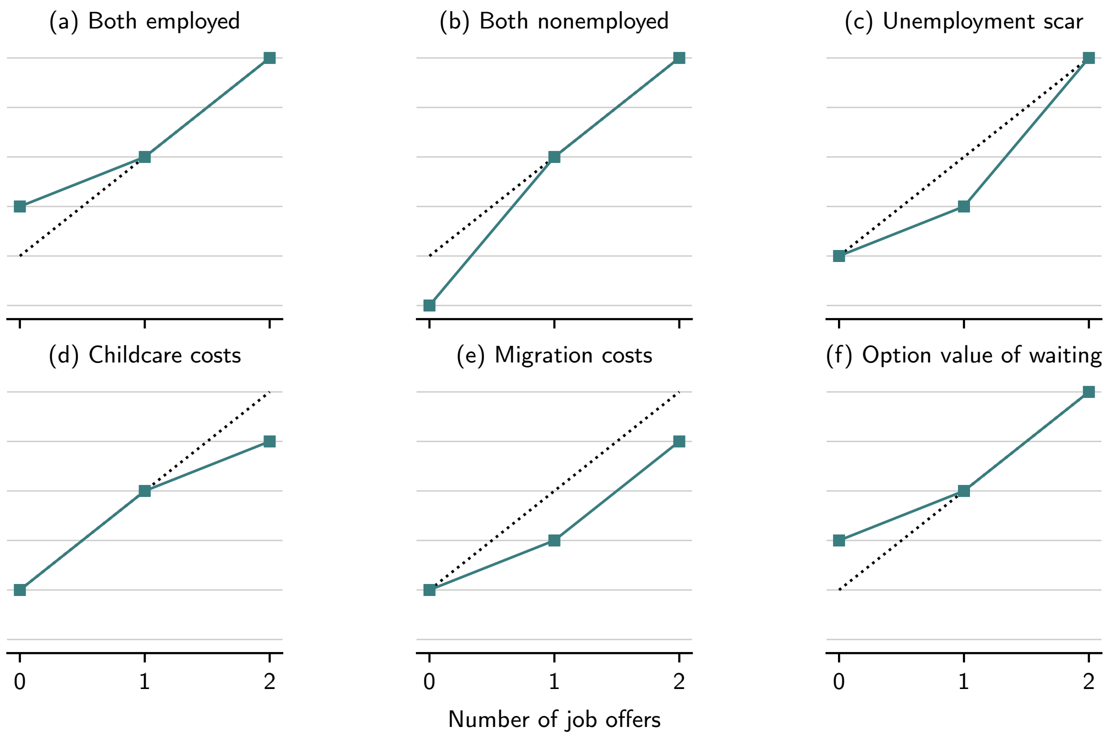

Figure 2: Illustration how distinct economic forces affect the curvature of $V - \chi$. The colocation friction is slack in the cases depicted in panels (b) and (d), and binds in the other cases.

where N is the number of commuting zones $q \neq r_0$.

Economic forces behind the colocation friction Evaluating condition (9) for the simple example, the colocation friction binds if and only if

$$\frac{u_0 + \Omega(N^{\frac{1}{1+\eta}})}{2} + \frac{u_2 - \rho\chi}{2} \geq u_1 - \rho\chi \tag{10}$$

with

$$\Omega(x) = 2\bar{\kappa}\left(\frac{x-1}{\rho + 2\bar{\kappa}x}\right)\left(u_1 - u_0 - \rho\chi + \frac{\bar{\kappa}}{\rho + \bar{\kappa}}(u_2 - u_1)\right).$$

The condition reveals several distinct factors affecting the curvature of $V - \chi$, illustrated in Figure 2. First, consider the case where there are no migration costs, $\chi = 0$, and search effort is infinitely elastic across locations, $\eta \to \infty$. In this case, condition (10) simplifies to $(u_0 + u_2)/2 \geq u_1$ so the curvature of $V - \chi$ in the number of simultaneous job offers is entirely determined by the curvature of u.

In particular, given u_1 and u_2, V is more convex for larger u_0, reflecting that the colocation friction is more likely to bind for households with a high initial value flow such as when both spouses are employed (illustrated in Panel a of Figure 2). Conversely, the colocation friction is more likely to be slack for households with a low initial value flow such as when both spouses are non-employed (Panel b).

Similarly, given u_0 and u_2, V is more convex the smaller u_1 (Panel c). Intuitively, this captures, in reduced form, economies where falling off the job ladder comes with high costs; e.g., due to human capital depreciation (e.g., Jung & Kuhn 2019), deteriorated bargaining positions (e.g., Cahuc, Postel-Vinay & Robin 2006; Lise & Robin 2017), or slippery bottom job rungs (Jarosch 2022). All of these mechanism increase the cost of becoming a trailing spouse, making it more likely that the colocation friction binds. Conversely, factors reducing the value flow of dual employment u_2, such as childcare costs, reduce the convexity of V and make it less likely that the colocation friction binds (Panel d).

Next, consider the case where migration costs are strictly positive, $\chi > 0$. As illustrated in Panel (e), this trivially introduces convexity in the search gains, making it more likely that the colocation friction binds. Intuitively, this is because migration costs accrue regardless of whether the household has one or two job offers at hand, which raises the relative returns of having two job offers.

Finally, consider the case where the search elasticity across locations η is finite (Panel f). In this case, there is an option value of delaying migration reflected in $\Omega(N^{\frac{1}{1+\eta}}) > 0$. Intuitively, with $\eta < \infty$, there are efficiency gains from broadening search to multiple locations. After migrating, these efficiency gains are lost when the trailing spouse narrows their search to the commuting zone of the leading spouse.[5] Again, this introduces convexity and makes it more likely that the colocation friction binds.

4 Quantitative Model

We now introduce the quantitative model and calibrate it to the U.S. labor market. The quantitative model is a special case of the general framework in Section 2, with locations corresponding to U.S. commuting zones and households being differentiated by several characteristics beyond their employment status and residence.

4.1 Setup

Household heterogeneity We quantify our model for dual-earner households whose adult members are composed of one woman, indexed by $i = f$, and one man, indexed by $i = m$. Households are differentiated by a vector $\boldsymbol{s} = (o_f, o_m, h_f, h_m, k)$, which along with households'

[5]While it is feasible to continue searching for jobs at multiple locations after migrating, this entails additional search and migration costs, which may or may not deter further migration. In either case, because search and migration costs are sunk, there is a positive option value of delaying migration if η is finite.

employment status and their location defines their type $(\boldsymbol{e}, \boldsymbol{s}, r)$. Specifically, each spouse i is characterized by an immutable occupation, o_i, and a time-varying human capital level, $h_i \in \{\underline{h}, \overline{h}\}$. In addition, we differentiate households with and without children, $k \in \{0, 1\}$.

Shocks Other than through search and migration, household types $(\boldsymbol{e}, \boldsymbol{s}, r)$ evolve stochastically through several independent shocks. For reasons that will become apparent below, we allow the arrival rates to vary across "occupation pairs", $\boldsymbol{o} \equiv (o_f, o_m)$, which constitute the immutable portion of the household type.

First, jobs are destroyed at an exogenous, gender-specific rate $\delta_i(\boldsymbol{o})$. Second, children arrive in and exit from households at rates rates $\pi_{k\uparrow}(\boldsymbol{o})$ and $\pi_{k\downarrow}(\boldsymbol{o})$. Third, human capital appreciates or depreciates as a function of a workers' current employment status: Employed workers' human capital appreciates to $h_i = \overline{h}$ at rate $\pi_{h\uparrow}(\boldsymbol{o})$, while nonemployed workers' human capital depreciates to $h_i = \underline{h}$ at rate $\pi_{h\downarrow}(\boldsymbol{o})$. Finally, households are exposed to location preference shocks (detailed below) that induce them to relocate from commuting zone r to q at rate $\pi_{q|r}(\boldsymbol{o})$.

Preferences and technology The instantaneous utility and labor product is given by

$$u(\boldsymbol{e}, \boldsymbol{s}, r) = \sum_{i \in \{f,m\}} \underbrace{\Big\{ h_i z_i(o_i, r) \cdot \mathbb{1}_{e_i=1} + b z_i(o_i, r) \cdot \mathbb{1}_{e_i=0} \Big\}}_{\text{labor product + home production}} + \underbrace{a(k, r) - p(r)}_{\text{amenities} - \text{rent}} - \underbrace{\xi(r) \cdot \mathbb{1}_{k=e_1=e_2=1}}_{\text{child care cost}}.$$

Here, $z_i(o_i, r)$ is a gender-, occupation- and commuting-zone-specific productivity that scales the labor and home products, $h_i z_i(o_i, r)$ and $b z_i(o_i, r)$. The value of living in commuting zone r is further determined by the value of its amenities $a(k, r)$ which differs by child status k, its cost of living $p(r)$, and its childcare costs $\xi(r)$. The latter accrue only if a household has children and both spouses are employed.

4.2 Calibration

We calibrate the model using household data from the American Community Survey (ACS) and the Current Population Survey (CPS), which we combine with geolocation data, commuting zone data from the Opportunity Atlas (Chetty, Friedman, Hendren, Jones & Porter 2018) and our own web-scraped data that inform the geography of the U.S. labor market. Appendix B.1 describes the data sources in detail.

Geography We consider all commuting zones (CZ) in the 48 contiguous states, excluding rural Illinois, Florida, and a few other isolated commuting zones for which we lack the data to construct our amenity measure.

For computational efficiency, we merge commuting zones that are close geographically and in terms of all observational characteristics. We describe the details of our recursive merging algorithm in Appendix B.4. Starting from 690 commuting zones with non-missing data, we obtain a final data set of 517 commuting zones post-merge.

Location preference shocks We design the location preference shocks, summarized by their arrival rates $\pi_{q|r}(\boldsymbol{o})$, to match the cross-sectional distribution of occupation pairs $\boldsymbol{o}$ over commuting zones. To do so in a computationally feasible way, we assume that location shocks induce a lump-sum utility shift,

$$\tau(\boldsymbol{e}, \boldsymbol{s}, r, q) = -\Big(V(\boldsymbol{e}, \boldsymbol{s}, q) - \chi(q|\boldsymbol{s}, r) - V(\boldsymbol{e}, \boldsymbol{s}, r)\Big),$$

that makes an household exactly indifferent to migrate from q to r.[6] With this design, households' value function and search policies, characterized by (4), are independent of the distribution of location shocks $\{\pi_{q|r}(\boldsymbol{o})\}$. For a given model parameterization, we can thus solve for V and the endogenous search and migration policies in a first step, and then design $\pi_{q|r}(\boldsymbol{o})$ in a second step to match, at the steady state, the empirical distribution over commuting zones.

We do so by choosing the distribution of shocks $\{\pi_{q|r}(\boldsymbol{o})\}$ with the smallest total prevalence that still allows us to exactly match the empirical distribution over commuting zones and occupations pairs $\boldsymbol{o}$. Our algorithm does so in a computationally efficient way: On an Intel i7-9700 computer, solving for V, identifying $\pi_{q|r}(\boldsymbol{o})$, and solving for occupation pair $\boldsymbol{o}$'s steady-state distribution takes about 2.5 seconds.

Assigned parameters We parameterize the model at a monthly frequency. Households retire and are replaced by new households at a monthly rate of 0.021/12, chosen so that the average households' work life lasts for 47 years. The effective discount rate ρ is set to the retirement rate plus 0.05/12, corresponding to an annual time preference rate of 5%.

We categorize households using the occupation classification by Autor & Dorn (2013). To economize on states, we drop occupations that are pursued by less than 3% of workers per gender.[7] This yields 18 occupation pairs, $(o_f, o_m) \in \{1, \ldots, 3\} \times \{1, \ldots, 6\}$, that cover 93.5% of all dual earner households in the ACS. Table 1 displays their frequency distribution.

[6] To avoid indirect effects on the employment distribution, we assume that location shocks to not alter the employment status $\boldsymbol{e}$.

[7] Following this criterion, we drop three occupations for women ("precision production and crafts", "machine operators assemblers and inspectors", and "transportation/ construction/ mechanics agricultural occupations") and do not drop any occupations for men.

Table 1: Frequency distribution over household occupation pairs

	o_m						
o_f	1	2	3	4	5	6	total
1	.32	.06	.03	.02	.02	.10	.55
2	.11	.04	.02	.01	.01	.09	.28
3	.04	.02	.03	.01	.01	.06	.17
total	.47	.12	.08	.04	.04	.25	1.00

Notes.–Frequencies are computed among dual-earner households in the ACS where o_f is the woman's occupation and o_m is the man's. Occupation codes follow the classification by Autor & Dorn (2013): (1) management/professional/technical/financial sales/public security, (2) administrative support and retail sales, (3) low-skill services, (4) precision production and crafts, (5) machine operators, assemblers and inspectors, (6) transportation/construction/mechanics/mining/agricultural occupations.

We set the job separation rates $\delta_i(\boldsymbol{o})$ based on the gender $\times$ occupation-specific employment to non-employment rates in the CPS. We use a Cobb-Douglas matching function, $p(\theta) = \theta^\gamma$, with matching elasticity $\gamma = 0.2$ as estimated by Lange & Papageorgiou (2020). We set the search elasticity across locations to $\eta = 0$ so search effort is inelastic across submarkets as in Albrecht et al. (2006), Kircher (2009) and Galenianos & Kircher (2009). For computational efficiency, we limit workers to a maximum of four simultaneous searches and verify ex post that the restriction is locally non-binding almost everywhere.[8]

Next, we normalize the high human capital realization to $\overline{h} = 1$ and set $b = 0.4$ so the home product equals 40% of a high-h workers' labor product (c.f., Shimer 2005). With b pinned down, we set $\underline{h} = b = 0.4$, capturing the idea that low-h workers are marginally attached to the labor market. Whether low-h workers join the labor force then depends on the value gain from future human capital appreciation versus the childcare costs that incur if they have children and their spouse is employed.

Our calibration of local productivities, $z_i(o_i, r)$, exploits that for almost all (i, o_i, r) the median worker in our model has high human capital.[9] With this is mind, we infer the local productivities $z_i(o_i, r)$ from the gender $\times$ occupation-specific median wages in a commuting zone r, as measured in the ACS.[10] We set their scale to normalize economy-wide average

[8]We verify this in the calibrated model by increasing the maximum number of simultaneous searches to five. We find that less than 0.001% of workers search in more than four markets simultaneously.

[9]Intuitively, the median worker is determined by the appreciation and depreciation rates $\pi_{h\uparrow}$ and $\pi_{h\downarrow}$, which we estimate below to match the *average* unemployment scar in the data. Given $\overline{h} - \underline{h}$, which is about five times the size of the average impact scar, our estimated process implies that the majority of workers has "high" (or, more accurately, "median") human capital. We verify numerical that conditional on (i, o_i, r), the median workers are indeed of type $\overline{h}$ for almost all gender $\times$ occupation $\times$ commuting zone combinations.

[10]Equating productivities with wages is model-consistent for two separate reasons. First, for labor contracts to be self-enforcing in the absence of contractual commitments on the worker-side, workers must be paid their labor product at all times other than during the instant they are hired (see, e.g., Menzio & Shi 2011; Baley, Figueiredo & Ulbricht 2022). Second, while our calibration uniquely pins down the vacancy cost c relative to search budgets $\bar{\kappa}_i(\boldsymbol{e}, \boldsymbol{s})^{(1-\gamma)/\gamma}$, their absolute value is indeterminate. Without loss of generality, we can

earnings to 1.

Next, we use average rents for two bedroom apartments to inform the cost of living $p(r)$ for each commuting zones. We set childcare costs $\xi(r)$ to 8.5% of the median household income in each commuting zone, consistent with the average, age-weighted household expenditures on childcare documented by Guner, Kaygusuz & Ventura (2020). We set the monthly arrival rate of children to the fertility rate among childless households in the ACS, $\pi_{k\uparrow} = 0.075/12$. We then use the departure rate of children to match the share of households with children in the ACS, which is 56%, yielding $\pi_{k\downarrow} = (0.56^{-1} - 1) \cdot \pi_{k\uparrow}$ net of the retirement rate.[11]

Finally, we combine several existing and web-scraped data sources to inform the local amenity values $a(k, r)$. Our data include information on crime rates, various climate and weather categories, walkability scores, measures of beach access and quality, and various data on infrastructure such as school and hospital quality or local government expenditures. We assume that local school quality is valued by households with children, while all other amenities are valued by all households. Appendix B.3 describes the data in detail.

One challenge in calibrating the amenity value is that $a(k, r)$ enters $u(\boldsymbol{e}, \boldsymbol{s}, r)$ in *income-equivalent* units, requiring us to convert the various data into income-equivalent units. To do so, we assume that the income-equivalent amenity value has the same passthrough rate on local rents as cross-regional differences in wages. Given the assumption, we can then infer the amenity value from the following regression:

$$p_r = \beta_0 + \epsilon \cdot \left(\overline{w}_r + \beta_1' \boldsymbol{a}_r^{\text{all}} + \beta_2' \boldsymbol{a}_r^{\text{kids}}\right) + v_r, \tag{11}$$

where p_r are average rents in commuting zone r, $\overline{w}_r$ is the local median household income, and $\boldsymbol{a}_r^{\text{all}}$ and $\boldsymbol{a}_r^{\text{kids}}$ collect our various amenity data (applicable to all households and only for households with children, respectively). We estimate a passthrough rate ϵ of 0.204, which is consistent with there being significant mobility frictions (as our framework indeed delivers). Given our estimate for ϵ, we infer the local amenity values as

$$a(k, r) = a_0 + \beta_1' \boldsymbol{a}_r^{\text{all}} + \beta_2' \boldsymbol{a}_r^{\text{kids}} \cdot \mathbb{1}_{k=1},$$

for some constant a_0. Without loss of generality, we normalize a_0 so that $\min_r a(0, r) = 0$.[12] Figure 3 plots the estimated amenity values for $k = 0$.

We summarize all exogenously assigned parameters in Table 2.

thus consider the limit where both $c \to 0$ and $\bar{\kappa}_i(\boldsymbol{e}, \boldsymbol{s}) \to 0$, in which case workers are always paid their labor product under the unique equilibrium labor contract.

[11]To be consistent with our "perpetual youth" setting, we re-weigh the ACS sample when computing the fertility rate and the share of households with children so that age is distributed geometrically in the sample.

[12]Constant shifts in $a(k, r)$ translate to constant shifts in V by a_0/ρ and are thus of no consequence.

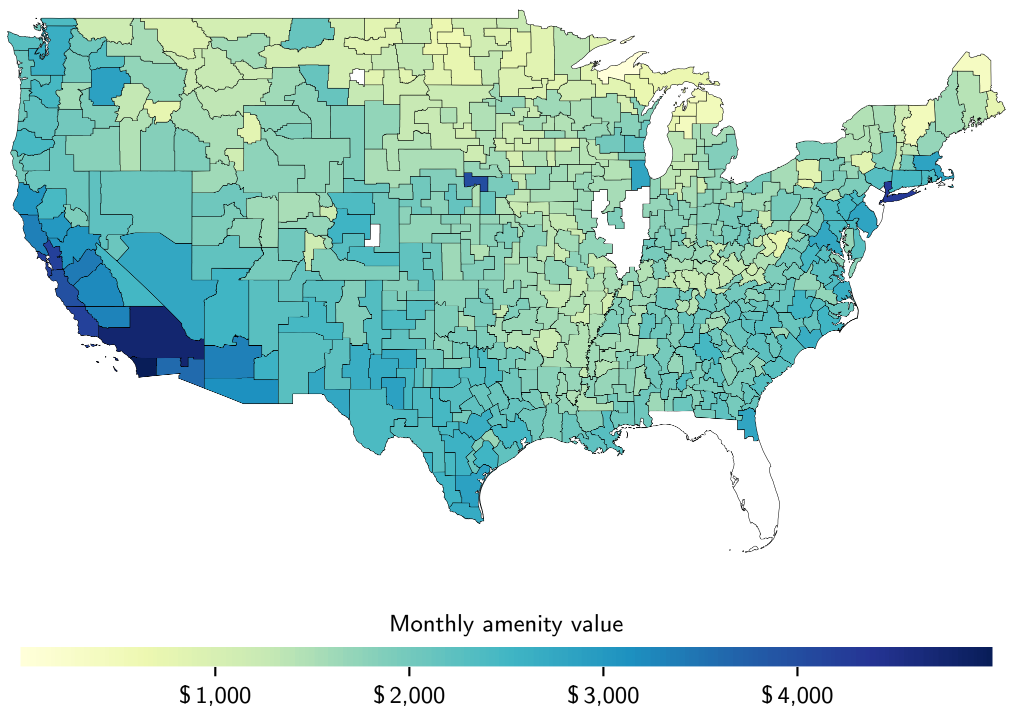

Figure 3: Income-equivalent amenity value. The plot shows the monthly amenity value for households without children ($k = 0$). The range of amenity values is normalized so that $\min_r a(0,r) = 0$. Commuting zones with missing data are filled in white.

Table 2: Assigned parameters

Parameter	Value	Source
Time preference rate, annualized	.05	literature
Retirement rate, annualized	.021	avg. working life of 47 years
Home production, $\bar{b}$	.4	Shimer (2005)
High human capital, $\bar{h}$	1.0	normalization
Low human capital, $\underline{h}$	.4	same as $\bar{b}$, see text
Search elasticity across locations, η	.0	literature
Matching elasticity, γ	.2	Lange & Papageorgiou (2020)
Child arrival rate $\pi_{k\uparrow}$, annualized	.075	ACS
Child departure rate $\pi_{k\downarrow}$, annualized	.038	ACS
Job separation rates, $\{\delta_i(\boldsymbol{o})\}$	see text	CPS
Local productivities, $\{z_i(o_i, r)\}$	see text	ACS
Cost of living, $\{p(r)\}$	see text	Opportunity Atlas
Child care costs, $\xi(r)$	see text	Guner, Kaygusuz & Ventura (2020)
Amenities, $\{a(k, r)\}$	see text	Opportunity Atlas, web scraped

Estimated parameters We calibrate the remaining parameters using the method of moments, with weights chosen to minimize the relative distance between model and empirical moments. All model moments are computed at the steady state. Our estimation leverages that the steady state distribution is generated by 18 non-communicating Markov chains, separated by occupation pairs $\boldsymbol{o}$. In choosing target moments and parameters that are also indexed by $\boldsymbol{o}$, we can therefore split the estimation into 18 subproblems, greatly reducing the computational complexity. As usual, conditional on $\boldsymbol{o}$, all parameters are identified jointly. In the following, we provide a heuristic mapping from moments to parameters to guide intuition.

In our model, the strength of search frictions is determined by the search budgets relative to the vacancy cost, $\bar{\kappa}_i(\boldsymbol{e}, \boldsymbol{s})/c^{\gamma/(1-\gamma)}$, which aren't separately identified. For some arbitrary normalization of c, we parameterize $\bar{\kappa}_i(\boldsymbol{e}, \boldsymbol{s}) = \bar{\kappa}_i(\boldsymbol{o})$, and use it to match the job finding rate out of unemployment at the gender $\times$ occupation-pair level, as computed in the CPS. To be consistent with the data, we only consider non-employed workers that are actively searching for jobs when computing the job finding rate in the model.

Next, following, e.g., Jung & Kuhn (2019) and Jarosch (2022), we estimate the human capital appreciation and depreciation rates, $\pi_{h\uparrow|e}(\boldsymbol{o})$ and $\pi_{h\downarrow|u}(\boldsymbol{o})$, to match the empirical steepness of job ladders. To do so, we simulate the wage scar of male workers separated from their job at $t = 0$ relative to the control group of non-separated workers, $\log(w_t^{\text{treat}}/w_t^{\text{control}})$, and match it to the estimate wage scars in Huckfeldt (2022) for $t = 12$ and $t = 36$ months.

It remains to calibrate the migration costs $\chi(q|\boldsymbol{s}, r)$. To do so, we differentiate between "work-related" migration, which in the model corresponds to the endogenous migration through job search, and other "residual" migration, captured in the model through location shocks. On the empirical side, we consider 45% of all observed migration as work-related, based on the survey evidence in Maurer (2017).

With this distinction, we calibrate migration costs by targeting, for each occupation pair $\boldsymbol{o}$, both the rate of work-related migration and its distribution in the ACS. Specifically, most migration in the data occurs at short distances and between commuting zones with similar population sizes. To capture these facts, we parameterize $\chi(q|\boldsymbol{s}, r)$ in terms of the spatial distance between any two commuting zones' population-weighted centroids, $d^{\text{geo}}(r, q)$, and the absolute difference in their population sizes, $d^{\text{pop}}(r, q)$. Specifically, we set

$$\chi(q|\boldsymbol{o}, r) = \sum_{k \in \{\text{geo}, \text{pop}\}} \sum_{j \in \{1, \ldots, 4\}} \chi_j^k(\boldsymbol{o}) \cdot \mathbb{1}_{d^k(r,q) \in \text{Bin}_j^k},$$

with four geographic and four population bins, $\{\text{Bin}_j^{\text{geo}}, \text{Bin}_j^{\text{pop}}\}_{j=1}^4$, chosen to capture the cross-bin variation in migration rates (c.f. Figure 5). For each occupation pair $\boldsymbol{o}$, we normalize $\chi_1^{\text{geo}}(\boldsymbol{o}) = 0$ and then estimate the seven remaining cost parameters, $\{\chi_j^k(\boldsymbol{o})\}$, to match the

Table 3: Summary of target moments

Moment	Level	No. of moments	Source	$\approx$ maps to
Job finding rate	$o_1 \times o_2 \times$ gender	36	CPS	$\bar{\kappa}_i(\boldsymbol{o})$
Wage scar, 1 & 3 yrs	$o_1 \times o_2 \times$ year	36	Huckfeldt (2022)	$\pi_{h\uparrow}(\boldsymbol{o})$, $\pi_{h\downarrow}(\boldsymbol{o})$
Migration rate, by bin	$o_1 \times o_2 \times$ bin	144	ACS	$\chi(q\|\boldsymbol{o}, r)$
Distribution over CZs	$o_1 \times o_2 \times r$	9306	ACS	$\pi_{q\|r}(\boldsymbol{o})$

Notes.–Due to adding up constraints on the distributions, the number of linearly independent moments is reduced by 36. The number of independent moments is equal to the number of estimated parameters.

Table 4: Untargeted moments

Moment	Model	Data
Employment rate, women	0.85	0.68
Employment rate, men	0.93	0.85
Dual earner share	0.80	0.61
Gender wage ratio, female-to-male	0.64	0.64
Gender earnings ratio, female-to-male	0.59	0.51
Residual migration, annualized ($\times$100)	1.28	1.39

total work-based migration rate of occupation pair $\boldsymbol{o}$ as well as its frequency distribution over the four geographic and the four population bins.

Table 3 summarizes the moments targeted in our estimation.

4.3 Model Validation

Targeted moments By design, our model matches *exactly* the cross-sectional distribution of households over commuting zones by occupation pairs. Figures 4 and 5 compare the remaining 216 moments with their data targets. The model fits these moments almost perfectly as well.

Untargeted moments Table 4 exhibits the model's fit vis-à-vis few non-targeted moments that are relevant in our context.

First, consider the employment rates by gender. While we target job-finding rates out of *unemployment*, employment rates crucially depend on a labor participation choice. As explained above, the labor participation choice in turn depends on the fraction of low-h workers, their productivity at home relative to the market, their option value of gaining work experience, and childcare costs – all of which are calibrated independently of the employment rate. The estimated employment rates are somewhat too large in the model, especially for women, suggesting that our calibration somewhat understates the relative value of non-employment. One possible omission that may explain this discrepancy is the absence of cultural norms in our model that are known to reduce female labor market participation

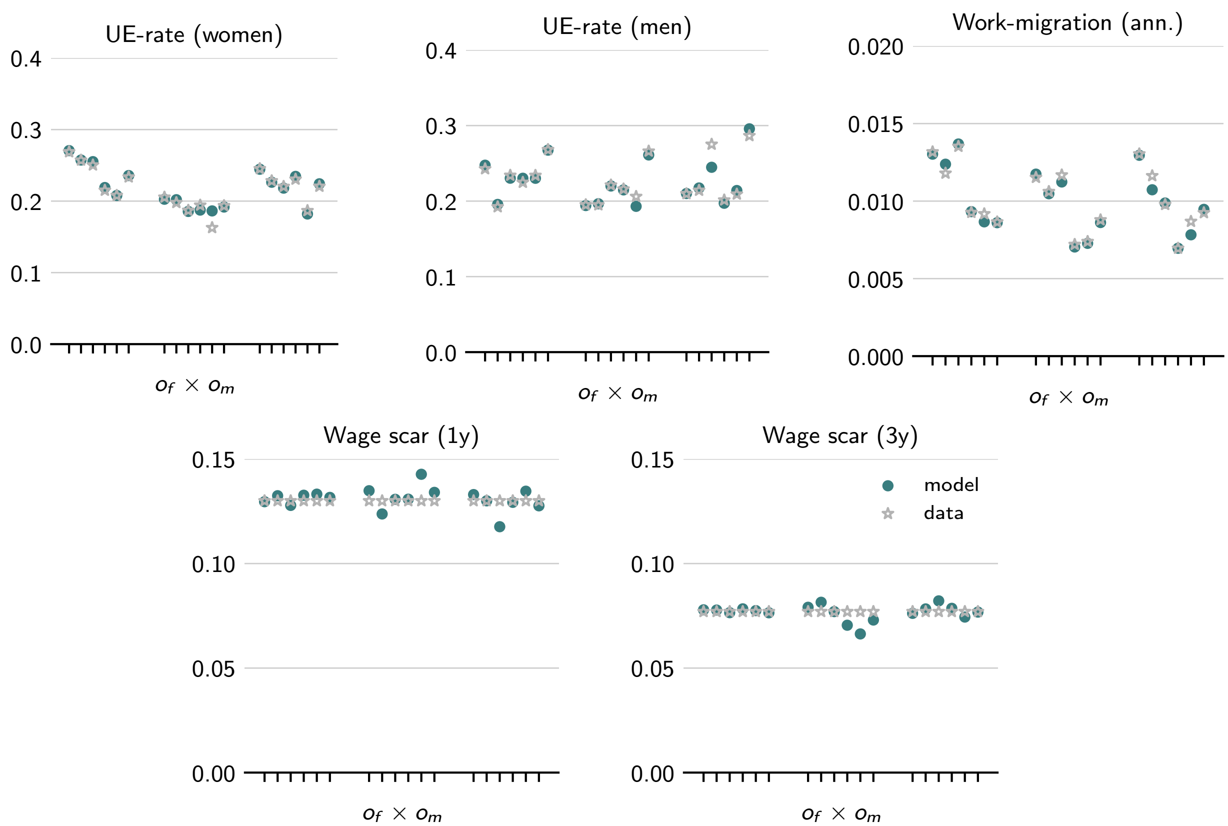

Figure 4: Targeted moments: Job finding rates, annualized migration rates, and wage scars by occupation pairs. Occupation pairs are grouped by the woman's occupation, o_f, with each individual tick corresponding to the man's occupation, o_m.

(Fernández 2013; Fernández & Fogli 2009).

Second, with regards to the gender wage distribution, the model captures both the wage and earning gaps in the data. In addition to the median wages (which we target), the gender wage distribution notably depends on the cross-sectional and cross-gender distributions of human capital, which are untargeted.

Third, it is worth briefly discussing the prevalence of migration due to location preference shocks. Our estimation targets 45% of the empirical migration to be explained by work-based migration, while minimizing the residual migration from location shocks subject to matching the cross-sectional distribution over commuting zones. As a sanity check, it is worthwhile to verify that the frequency of location shocks does not exceed the residual migration in the data. Reassuringly, this is indeed the case.

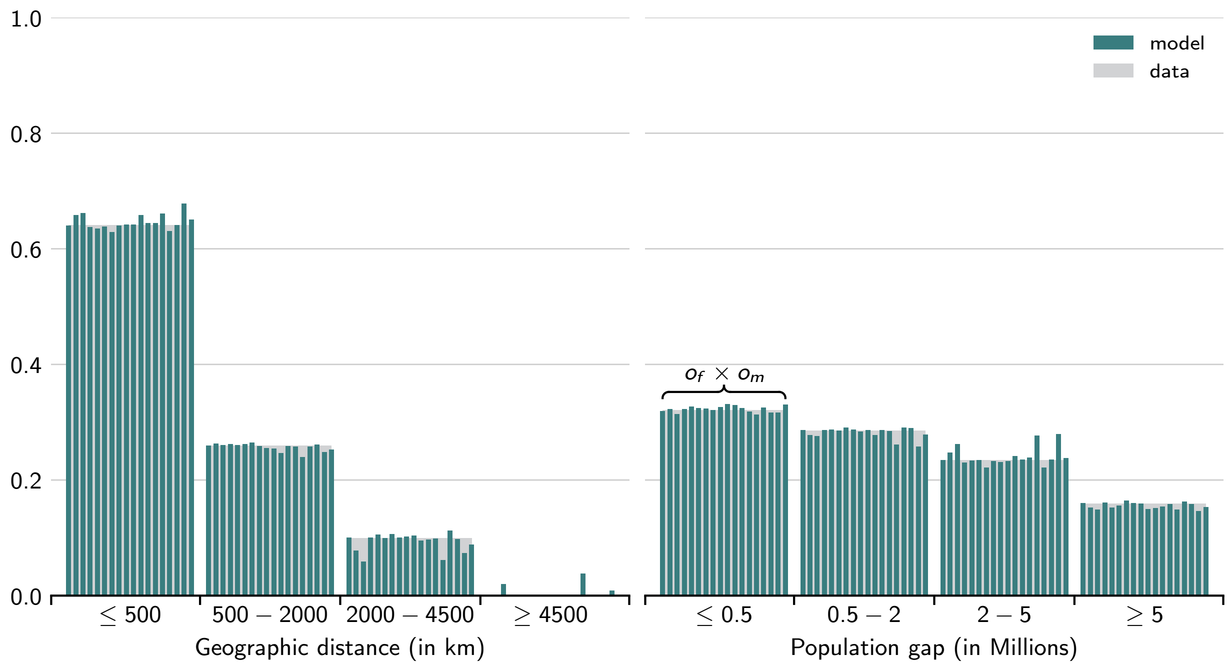

Figure 5: Targeted moments: Migration frequency over geographic distance and over population gap

5 Quantifying the Colocation Friction

We are now ready to study the consequences of the colocation friction for the U.S. labor market. Applying the characterization in Proposition 1 to the estimated model, we find that the colocation friction binds in virtually all non-local searches, $q \neq r$, and is slack in virtually all local searches, $q = r$. In the remainder of this section, we quantify the consequences of when the friction binds, beginning with the consequences for migrating couples.

5.1 Consequences Conditional on Migration

To provide context, consider first the labor market experience of migrating households in the baseline economy. About three quarter of all work-based migration is initiated by a job offer to men; so women are about three times as likely as men to become a trailing spouse.

We assess the implications of this imbalance by simulating the trajectory of a representative sample of households with a single migration event at $t = 0$, and compare it with the trajectory of the same sample of households absent any migration. The solid lines in Figure 6 show the log-differences between the two groups. Cumulatively over their lifetime, the average migrating man experiences earnings gains of 56.9% compared to the control group. These gains are mostly driven by differences in local productivities.

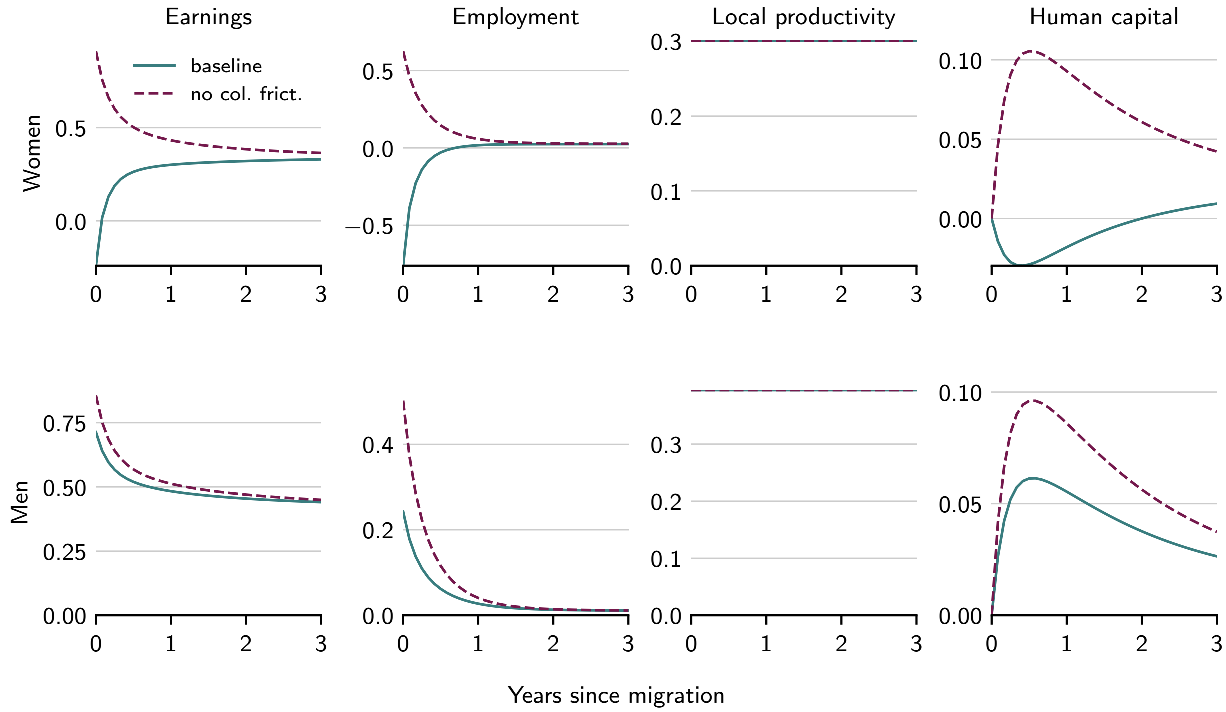

Figure 6: Consequences of migration. The plots show log-differences relative to the counterfactual of no migration. The difference between the baseline and the no-colocation-friction benchmark amounts to the direct effect of the colocation friction.

The migration experience of men contrasts starkly with that of women. In absolute terms, mens' migration gains are about 2.5 times that of women. In parts, this reflects the earnings gap across genders. Still, compared to their own control group, the average womens' migration gains (37.4%) are significant less than that of men (c.f. Table 5). On the one hand, this is because women are more likely to be jobless post migration, *reducing* their earnings and human capital in the first few months after migration. On the other hand, in the long run, women tend to benefit less from migration with an average increase in productivity of about 35% compared to 50% by men.

We next isolate the causal effect that the colocation friction has on the migration experience of households. To do so, we consider the same migrating sample of households, but now simulate their trajectory under the benchmark where the colocation friction is relaxed (dashed lines in Figure 6).[13] Not surprisingly, the benchmark is close to the baseline for men, reflecting that few men are trailing spouses. By contrast, we see huge short-term differences compared to the baseline for women, capturing both the immediate effect on employment and the indirect effect on human capital. While the employment effect is all but disappeared after 1

[13]Here we keep both the sample of households and their migration destinations fixed, changing only the initial employment status according to the optimal correlation policies under the benchmark.

Table 5: Consequences of migration on lifetime earnings

		ΔPV in earnings	
	Leading Spouse	Baseline economy	No colocation friction
Women	24.1 %	.241 (37.4 %)	.307 (47.7 %)
Men	75.9 %	.591 (56.9 %)	.616 (59.3 %)

Notes.—ΔPV in earnings denotes the difference in the present value of earnings between a representative sample of migrating households and the counterfactual without migration. Without parentheses are the gains in level, denominated relative to the economy-wide average value of lifetime earnings. In parenthesis are the same numbers relative to the present value of earnings in the gender-specific control group.

year, the indirect effect from collecting work experience persists for a few years. Cumulated over their lifetime, the colocation friction reduces the migration gains of women by 27% (or, equivalently, by 6.6% of the *lifetime* earnings of the average worker), with most of these losses accruing in the first 3 years.[14]

5.2 Discouraged Migration

As hypothesized by Mincer (1978), the colocation friction may not only affect the labor market experience for households that migrate but, perhaps more importantly, also affect households' propensity to migrate in the first place.

To quantify the relevance of Mincer's hypothesis, we next relax the colocation friction for a representative, zero-measure sample of households and their descendants, and study their migration behavior in the sequel.[15] Figure 7 plots the migration rate for these households. After the colocation friction is relaxed, work-migration increases initially by 50%, and then converges to its new steady state level which continues to be about 20% above the baseline level.[16]

Given the difference in migration rates, we next ask who are the households that are discouraged from migrating due to the colocation friction. To do so, Table 6 contrasts average characteristics of households that migrate when the colocation friction is active with those

[14] 90% of the losses accrue in the first 3 years, equal to 6.0% of the lifetime earnings of the average worker.

[15] As treated households retire, we replace them by an equal mass of newborns in order to ensure that the treatment sample converges to a steady state and is comparable to the baseline. If instead we do not replace retiring households, we find that migration rates in both the baseline and the treatment decline over time as the population ages. In this case, the log-gap between benchmark and baseline is the same at $t = 0$, but then *widens* over time compared to the one shown in Figure 7.

[16] Note that by relaxing the friction perpetually, momentary migration incentives are reduced by the option value of migrating without friction in the future. Alternatively, we can measure the "compound effect" of the colocation friction on momentary migration rates by relaxing the friction during only a short time window and then reimpose it after. In this case, migration more than triples to an annual rate of 3.65% while the friction is relaxed.

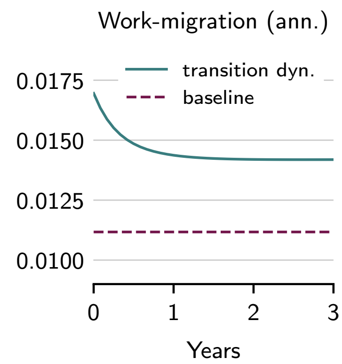

Figure 7: Impact of the colocation friction on work-migration. The dashed line shows the annual steady-state migration rate in the baseline economy. The solid line shows the rate for a representative sample of households for whom the friction is relaxed from $t = 0$ onward.

that only migrate when the friction is relaxed.

When the friction is active, migrating households tend to be childless with below-average household incomes. In line with the stylized examples in Section 3.3, these are precisely households for whom the colocation friction is least consequential. Crucially, their poverty tends to be caused by outside factors (employment status and local productivities) rather than their intrinsic earnings potential defined by their human capitals, so that migrating is indeed profitable.

By contrast, households who the friction discourages from migrating tend to be dual-employed with above-average incomes and extremely high human capital, resembling the stereotypical "power couple". Given their human capital, they have the highest potential returns from migrating, yet they are also most affected by the colocation friction given their above-average incomes and the looming loss of work experience for the trailing spouse (both of which favors a more convex shape of the value function as demonstrated in Section 3.3).

We quantify the economic consequences of discouraged migration, by computing the lifetime earnings losses posed by the friction on precisely those households that are discouraged from migrating. Measured again in units of average lifetime earnings, the earnings loss from the discouraged migration is 0.112 for women and 0.135 for men (or, equivalently, 13% and 10.5% of their own lifetime earnings subject to the friction).

5.3 Long-run Impact on Employment and Earnings

To quantify the long-run consequences of the colocation friction, consider again the sample of households for whom we have relaxed the colocation friction at $t = 0$. Figure 8 plots transition dynamics in their earnings, employment, local productivity, and human capital. Note that these effects are smaller by about an order of magnitude compared to the conditional dynamics

Table 6: Comparison of baseline migrants to discouraged migrants

	Migrating	Discouraged	Population average
Children	.19	.36	.56
Both employed	.34	.82	.80
Both nonemployed	.23	.01	.01
Household income	1.12	2.12	2.00
Employment rate			
women	.52	.89	.88
men	.60	.93	.93
Human capital			
women	.76	1.00	.79
men	.88	.99	.86
Local productivity			
women	.91	1.01	.98
men	1.29	1.34	1.50

Notes.–Displayed are averages for three household groups: (i) Migrating: households who migrate in the baseline economy. (ii) Discouraged: households who do not migrate in the baseline economy, but migrate when the colocation friction is relaxed. (iii) Population average: All households, weighted by the steady-state distribution.

in Figure 6, reflecting that even absent the friction, migration along the transition path occurs only at an annual rate of about 1.5%. In addition to the direct effect on migrating households seen in the conditional dynamics above, the unconditional dynamics further encompass the impact of the friction on discouraged migration, as well as adjustments in households' labor market choices in response to the friction.

Taking into account all these effects, there is a strong expansion in female employment immediately after relaxing the friction, raising female earnings by 0.76%. In the medium term, the additional work experience translates to human capital gains that raise female earnings by an additional 0.87%. Finally, in the long-term, continued relocation to more productive

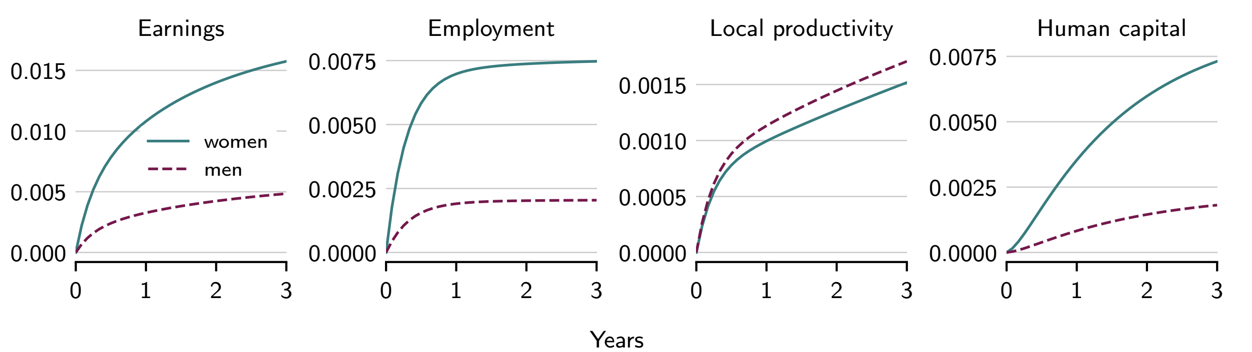

Figure 8: Impact of the colocation friction on earnings, employment, local productivities, and human capital. The plots show log-differences between a representative sample of households for whom the friction is relaxed from $t = 0$ onward and the steady state where the friction is active.

Table 7: Steady state comparison: baseline economy vs. correlated matching benchmark

	Women	Men
Earnings	2.07 %	.84 %
Employment	.76 %	.19 %
Local productivity	.47 %	.51 %
Human capital	.87 %	.20 %

Notes.—Differences are computed between the steady-state distribution of the baseline economy and the steady-state distribution of the benchmark without colocation friction.

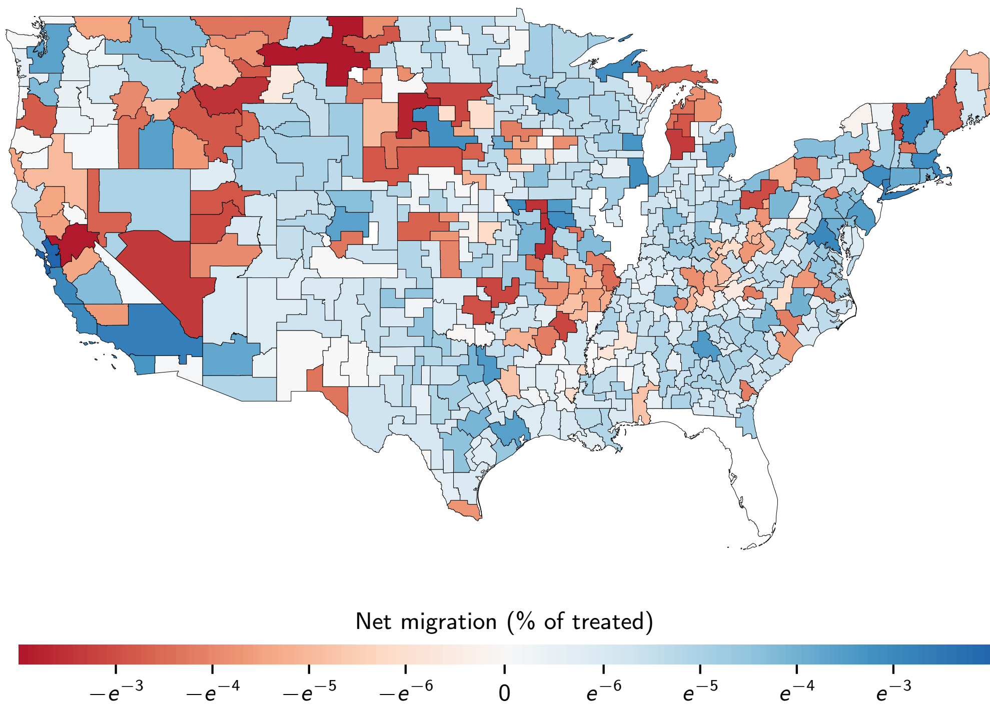

Figure 9: Steady-state distribution, baseline economy vs. measure 0 of households without colocation friction

commuting zones raises female earnings by another 0.47%. By contrast, the combined effect of the colocation friction on male earnings is more subdued.

Table 7 compared the long-run earnings differences across steady states. Taken together, the colocation friction reduces average steady-state earnings of women by 2.07% and reduces average steady-state earnings of men by 0.84%.

5.4 Long-run Impact on Location Choice

Next, we explore how the colocation friction impacts where households reside. At its new steady state, the distribution over commuting zones among households without the colocation friction diverges from the original steady-state distribution by 2.7%. Figure 9 plots the

Table 8: Steady state comparison: consequence of average relocation

Local productivities, women	.172
Local productivities, men	.287
Amenities	.285
Rents	.089
Childcare costs	.024

Notes.—Productivities are scaled to normalize average earnings to 1 (or, equivalently, average household income to 2), setting the units for all nominal quantities.

Table 9: Lifetime utility gains by household occupation pairs

	o_2						
o_1	1	2	3	4	5	6	mean
1	2.69	1.86	1.33	2.13	1.61	1.48	2.25
2	1.18	1.13	.91	.96	.78	.85	1.02
3	.64	.67	.37	.55	.48	.48	.53
mean	2.15	1.44	.88	1.38	1.03	1.03	1.61

Notes.–Gains are denominated in lifetime earnings equivalents and are denoted in percent. Here, o_1 is the woman's occupation and o_2 is the man's. Occupation codes follow the classification by Autor & Dorn (2013): (1) management/professional/technical/financial sales/public security, (2) administrative support and retail sales, (3) low-skill services, (4) precision production and crafts, (5) machine operators, assemblers and inspectors, (6) transportation/construction/mechanics/mining/agricultural occupations.

difference in steady-state distributions over commuting zones. Broadly, without the colocation friction, households would move from the Rocky Mountains and Midwest to the Pacific, Northeast, and South.

As summarized in Table 8 , these relocations would raise household earnings and amenities, while also raising the cost-of-living. For households with children that realize their earnings potential, $z_f + z_m$,[17] forgoing these relocations in light of the colocation friction imposes an income-equivalent loss of 0.63 or, equivalently, 32% of the average household income.

5.5 Consequences for Life-time Utility

In sum, the colocation friction adversely affects trailing spouse's careers, discourages migration, and affects where households end up living. To summarize the impact of these effects on overall welfare we contrast the lifetime utility of a household in the baseline economy with that of a household born without colocation frictions. We find that, in income-equivalent units, the loss in lifetime utility is equivalent to a 1.61% loss in lifetime earnings. Table 9 decomposes the gains by immutable occupation pairs. The gains are largest for households where both spouses work in "management, professional, technical, financial sales or public security professions", and are lowest for households where both spouses work in "low-skilled

[17]A worker realizes their earnings potential when $h_i = e_i = 1$, in which case their earnings $z_i h_i e_i$ equal z_i.

services".

6 Concluding Remarks

This paper develops a spatial directed search model that captures the unique frictions that characterize the job search by dual-earner households. We estimate the model for the U.S. labor market. The estimated model matches, at the occupation-level, both labor market and migration flows as well as the cross-sectional density of households over commuting zones. We find that dual-earner households are exposed to a "colocation-friction" that vastly reduces their gains from migration with long-run consequences for average employment and earnings in the economy. All in all, we estimate that the colocation friction incurs a lifetime utility loss equivalent to a 1.61% decrease in lifetime earnings.

Our framework is among the first that incorporates job search by dual-earner households into a spatial model of the labor market. It is distinguished from the existing literature by its analytical tractability, which opens the door to a large-scale estimation at the commuting-zone level. It is also the first framework that models dual-earner job search as directed, accounting for spouses ability to coordinate *search effort* across locations, thus exposing spatially mismatched *job offers* as the true underlying friction.

The framework delivers rich predictions regarding dual-earner households' careers and gender discrepancies in employment and earnings. We view future applications of our framework that further explore these aspects as well as the transformation to work-from-home jobs that come with fewer spatial constraints as fruitful avenues.

References

Albrecht, James, Pieter A Gautier & Susan Vroman (2006). “Equilibrium Directed Search with Multiple Applications”. *Review of Economic Studies* 73 (4), pp. 869–891.

Autor, David H. & David Dorn (2013). “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market”. *American Economic Review* 103 (5), pp. 1553–97.

Baley, Isaac, Ana Figueiredo & Robert Ulbricht (2022). “Mismatch Cycles”. *Journal of Political Economy* 130 (11), pp. 2943–2984.

Cahuc, Pierre, Fabien Postel-Vinay & Jean-Marc Robin (2006). “Wage Bargaining with On-the-Job Search: Theory and Evidence”. *Econometrica* 74 (2), pp. 323–364.

Chetty, Raj, John N Friedman, Nathaniel Hendren, Maggie R Jones & Sonya R Porter (2018). *The opportunity atlas: Mapping the childhood roots of social mobility.* Tech. rep. National Bureau of Economic Research.

Chetty, Raj et al. (2016). “The Association Between Income and Life Expectancy in the United States, 2001-2014”. *JAMA* 315 (16), pp. 1750–1766.

Costa, Dora L. & Matthew E. Kahn (2000). “Power Couples: Changes in the Locational Choice of the College Educated, 1940–1990”. *The Quarterly Journal of Economics* 115 (4), pp. 1287–1315.

Dey, Matthew & Christopher Flinn (2008). “Household search and health insurance coverage”. *Journal of Econometrics* 145 (1). The use of econometrics in informing public policy makers, pp. 43–63.

Din, Alexander & Ron Wilson (2020). “Crosswalking ZIP Codes to Census Geographies: Geoprocessing the U.S. Department of Housing and Urban Development’s ZIP Code Crosswalk Files”. *Cityscape: A Journal of Policy Development and Research* 22 (1).

Fernández, Raquel (2013). “Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century”. *American Economic Review* 103 (1), pp. 472–500.

Fernández, Raquel & Alessandra Fogli (2009). “Culture: An Empirical Investigation of Beliefs, Work, and Fertility”. *American Economic Journal: Macroeconomics* 1 (1), pp. 146–77.

Findeisen, Sebastian, Sang Yoon (Tim) Lee, Tommaso Porzio & Wolfgang Dauth (2021). “Transforming Institutions: Labor Reallocation and Wage Growth in a Reunified Germany”. *Working Paper.*

Flabbi, Luca & James Mabli (2018). “Household Search or Individual Search: Does It Matter?” *Journal of Labor Economics* 36 (1), pp. 1–46.

Galenianos, Manolis & Philipp Kircher (2009). “Directed Search with Multiple Job Applications”. *Journal of Economic Theory* 144 (2), pp. 445–471.

Galenianos, Manolis, Philipp Kircher & Gábor Virág (2011). “Market Power and Efficiency in a Search Model”. *International Economic Review* 52 (1), pp. 85–103.

Gemici, Ahu (2007). “Family Migration and Labor Market Outcomes”.

Guler, Bulent, Fatih Guvenen & Giovanni L. Violante (2012). “Joint-Search Theory: New Opportunities and New Frictions”. *Journal of Monetary Economics* 59 (4), pp. 352–369.

Guner, Nezih, Remzi Kaygusuz & Gustavo Ventura (2020). “Child-Related Transfers, Household Labor Supply and Welfare”. *Working Paper.*

Herkenhoff, Kyle, Gordon Phillips & Ethan Cohen-Cole (2022). “How Credit Constraints Impact Job Finding Rates, Sorting & Aggregate Output”. *Working Paper.*

Hosios, Arthur J. (1990). “On the Efficiency of Matching and Related Models of Search and Unemployment”. *The Review of Economic Studies* 57 (2), p. 279.

Huckfeldt, Christopher (2022). “Understanding the Scarring Effect of Recessions”. *American Economic Review*, forthcoming.

Jarosch, Gregor (2022). “Searching for Job Security and the Consequences of Job Loss”. *Econometrica*, forthcoming.

Jovanovic, Boyan (1979). “Job Matching and the Theory of Turnover”. *Journal of Political Economy* 87 (5), pp. 972–990.

— (1984). “Matching, Turnover, and Unemployment”. *Journal of Political Economy* 92 (1), pp. 108–122.

Jung, Philip & Moritz Kuhn (2019). “Earnings Losses and Labor Mobility over the Life Cycle”. *Journal of the European Economic Association* 17 (3), pp. 678–724.

Kaplan, Greg & Sam Schulhofer-Wohl (2017). “Understanding the Long-Run Decline in Insterstate Migration”. *International Economic Review* 58 (1), pp. 57–94.

Kennan, John & James R. Walker (2011). “The Effect of Expected Income on Individual Migration Decisions”. *Econometrica* 79 (1), pp. 211–251.
Kircher, Philipp (2009). “Efficiency of Simultaneous Search”. *Journal of Political Economy* 117 (5), pp. 861–913.
Lange, Fabian & Theodore Papageorgiou (2020). *Beyond Cobb-Douglas: Flexibly Estimating Matching Functions with Unobserved Matching Efficiency.* w26972. Cambridge, MA: National Bureau of Economic Research.
Lise, Jeremy & Jean-Marc Robin (2017). “The Macrodynamics of Sorting between Workers and Firms”. *American Economic Review* 107 (4), pp. 1104–1135.
Maurer, Roy (2017). *Americans Most Often Move for Work.* SHRM. URL: https://www.shrm.org/resourcesandtools/hr-topics/talent-acquisition/pages/americans-most-often-move-work.aspx (visited on 09/16/2022).
Menzio, Guido & Shouyong Shi (2010). “Block Recursive Equilibria for Stochastic Models of Search on the Job”. *Journal of Economic Theory* 145 (4), pp. 1453–1494.
— (2011). “Efficient Search on the Job and the Business Cycle”. *Journal of Political Economy* 119 (3), pp. 468–510.
Menzio, Guido, Irina A. Telyukova & Ludo Visschers (2016). “Directed Search over the Life Cycle”. *Review of Economic Dynamics* 19, pp. 38–62.
Mincer, Jacob (1978). “Family Migration Decisions”. *Journal of Political Economy* 86 (5), pp. 749–773.
Pilossoph, Laura & Shu Lin Wee (2021). “Household Search and the Marital Wage Premium”. *American Economic Journal: Macroeconomics* 13 (4), pp. 55–109.
Piyapromdee, Suphanit (2020). “The Impact of Immigration on Wages, Internal Migration, and Welfare”. *The Review of Economic Studies* 88 (1), pp. 406–453.
Rabinovich, Stanislav & Ronald Wolthoff (2022). “Misallocation inefficiency in partially directed search”. *Journal of Economic Theory* 206, p. 105559.
Şahin, Ayşegül, Joseph Song, Giorgio Topa & Giovanni L. Violante (2014). “Mismatch Unemployment”. *American Economic Review* 104 (11), pp. 3529–3564.
Schaal, Edouard (2017). “Uncertainty and Unemployment”. *Econometrica* 85 (6), pp. 1675–1721.
Shimer, Robert (2005). “The Cyclical Behavior of Equilibrium Unemployment and Vacancies”. *American Economic Review* 95 (1), pp. 25–49.
Snyder, J.P. (1982). *Map Projections Used by the U.S. Geological Survey.* Geological Survey United States: Geological Survey bulletin. U.S. Government Printing Office.
Venator, Joanna (2021). “Dual-Earner Migration Decisions, Earnings, and Unemployment Insurance”.
Wright, Randall, Philipp Kircher, Benoît Julien & Veronica Guerrieri (2021). “Directed Search: A Guided Tour”. *Journal of Economic Literature* 59 (1), pp. 90–148.

A Mathematical Appendix

A.1 Proof of Proposition 1

For a given value function, search policies in the correlated matching benchmark solve

$$\max_{\{f_{i,q},\omega_q,y_{i,q}\}} \left\{ \sum_{i,q} \left(f_{i,q} - \omega_q\right) \Delta V_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r) \Big) + \sum_{q} \omega_q \Delta V_q^{\text{corr}}(\boldsymbol{e}, \boldsymbol{s}, r) - \sum_{i,q} f_{i,q} y_{i,q} \right\}$$

subject to (2), (3), (7) and (8). Observe that the objective is separable across q and is linear in $\{\omega_q\}$. Thus, for any q, ω_q is optimally set to one of the boundaries in (8), with the upper boundary being optimal if and only if

$$\Delta V_q^{\text{corr}}(\boldsymbol{e}, \boldsymbol{s}, r) \geq \sum_i \Delta V_{i,q}(\boldsymbol{e}, \boldsymbol{s}, r).$$

This proves Proposition 1.

A.2 Kolmogorov Forward Equation

Collect all job finding rates $\{f_{i,q}(e, s, r\}$ along with all endogenous separations into μ, let n denote the replacement process of retiring households by new ones, and define $\phi \equiv \mu + \lambda + n$. Then the cross-sectional distribution, $g_t(\boldsymbol{e}, \boldsymbol{s}, r)$, evolves according to the following differential equation:

$$\frac{\mathrm{d}g_t}{\mathrm{d}t}(\boldsymbol{e}', \boldsymbol{s}', r') = \sum_{\boldsymbol{e},\boldsymbol{s},r} \phi(\boldsymbol{e}', \boldsymbol{s}', r' | \boldsymbol{e}, \boldsymbol{s}, r)\, g_t(\boldsymbol{e}, \boldsymbol{s}, r). \tag{A.1}$$

From (A.1), we may obtain the steady state distribution by setting $\mathrm{d}g = 0$ subject to $\sum g = 1$.

In practice, when estimating the model, we embed the steady-state condition as a constraint into our moments-matching algorithm when computing the distribution of location preference shocks with the smallest total prevalence subject to matching the cross-sectional distribution of households over commuting zones.

B Data Appendix

B.1 Data Sources

This appendix describes the data sources that we use in calibrating our model.

American Community Survey (ACS) We construct our ACS sample using the years 2010–2019, restricting attention to individuals aged 18–64, living in the 48 contiguous states (i.e., excluding Alaska, Hawaii, and Puerto Rico), who are part of a couple (married or cohabiting). 1990 Census Occupation codes are mapped into the occupation classification by Autor & Dorn (2013) using their crosswalk. We deflate wages to 2010 USD using the CPI. To obtain median wages and occupation-shares by commuting zone (CZ), we adopt crosswalks by Autor & Dorn (2013) to aggregate data from Public Use Micro Areas (PUMAs) to the CZ-level.

Current Population Survey (CPS) We construct our CPS sample replicating the sample restrictions used for the ACS. 1990 Census Occupation codes are mapped into the occupation classification by Autor & Dorn (2013) using their crosswalk.

Opportunity Atlas We use CZ data from the Opportunity Atlas. The Opportunity Atlas draws on various U.S. data sources and aggregates them on the CZ level (see Chetty et al. 2018). For a detailed description of several variables that we draw from the Opportunity Atlas see also Chetty et al. (2016).

Web-scraped data We use web-scraping to obtain CZ-level information on local weather and climate conditions, crime rates, walkability scores, measures of beach access and quality, school and hospital quality and local government expenditures. Specifically, we scrape information published on bestplaces.net, usnews.com, walkscore.com, and watersgeo.epa.gov. The raw data are aggregated on the county and ZIP code level, respectively. To aggregate these data to CZs, we use crosswalks by Autor & Dorn (2013) and Din & Wilson (2020). For a description of all web-scraped variables and their sources see Table A.I in Appendix B.3.

Geolocation data To locate CZs in space we use county centroid geographic coordinate system (GCS) coordinates provided by simplemaps.com. We apply the Albers equal-area projection to transform GCS coordinates to Albers equal-area (AEA) coordinates.[A1] The advantage of using AEA coordinates is that they are cartesian; i.e., geographic distances can be expressed as euclidean "straight-line" distances, which allows us to adopt the crosswalk by Autor & Dorn (2013) to aggregate county coordinates to population weighted CZ centroids and further aggregate commuting zones as described in Appendix B.4.

[A1] We follow Snyder (1982) in using the AEA standard parallels 29.5º and 45.5º north.

B.2 Measuring Labor Market and Migration Flows

Labor Market Flows We measure labor market flows by using the rotating panel structure of the CPS. The CPS is a monthly survey conducted among a nationally representative sample of households. Leveraging matched individual records for consecutive months, we compute monthly transition frequencies between employment states, conditioning on sex and household occupations.

Migration Flows We measure migration flows using ACS data on where individuals reside when surveyed and their residence one year prior. In particular, the ACS records the Public Use Micro Area (PUMA) individuals reside in when surveyed and the Migration Public Use Micro Area (MIGPUMA) one year prior. MIGPUMAs are constructed from one or multiple PUMAs. We map MIGPUMAs into PUMAs using a crosswalk published by usa.ipums.org, and map PUMAs into CZs using crosswalks developed by Autor & Dorn (2013). In computing across-CZ migration rates, we adopt a conservative approach, counting as cross-CZ migration only a change in residency from a MIGPUMA to a PUMA that have zero overlap in the CZs they intersect with. To obtain spatial distances between migration origin and destination, we use spatial coordinates of population weighted CZ centroids. For differences in population size we use population counts as recorded in the ACS. We aggregate to the PUMA and MIGPUMA level, respectively, using the crosswalks mentioned above together with a crosswalk from counties to CZs also by Autor & Dorn (2013).

B.3 Measuring Amenities

This appendix describes the data inputs and results of regression 11, which we use to convert our CZ-level amenity data into income equivalent units. Table A.I describes the definition and data source of each amenity in our data. Table A.II summarizes the estimation results of regression (11). Almost all coefficient estimates are statistically significant at the 1% and all are significant at the 10% level. The adjusted R^2 of the regression is 92%.

B.4 Merging Commuting Zones

This appendix describes the algorithm by which we merge CZs that are close geographically and in terms of observed characteristics. Geographic distances between CZs are computed based on AEA coordinates of CZ centroids. We define closeness in terms of all observed CZ characteristics that feed into our calibration; i.e., rents, amenities with and without children (converted to income equivalent units as described in Section 4.2), the local population counts of occupation pair $\boldsymbol{o}$, and gender $\times$ occupation-specific median wages. As a joint measure of

Table A.I: Description of amenities data

Variable	Description	Data Source
Summer climate score	Index that captures temperature, precipitation, average number of sunny days, freezing days, extremely freezing days. Scale: 1–10, 10 being the best. Based on April–Sept data.	bestplaces.net
Winter climate score	Index that captures temperature, precipitation, average number of sunny days, freezing days, extremely freezing days. Scale: 1–10, 10 being the best. Based on Oct–March data.	bestplaces.net
Population density	Number of people per square mile.	Opportunity Atlas
Local government expenditures	Total local government expenditures per capita, USD per annum.	Opportunity Atlas
Crime rate	Annual per capita crime rate per million people.	Opportunity Atlas
Walkability score	Score based on availability of infrastructure and number of restaurants, bars and coffee shops within 5 minutes walking distance. More choices within a radius yields a higher score. Scale: 0–100, 100 being the highest walkability. Aggregated to CZ level from Zip Code Level.	walkscore.com
Beach access (total length in miles)	The EPA's measurement of beach access (official beaches, not just shoreline) in each county. Aggregated to CZ level from county level.	watersgeo.epa.gov
No. of top tier beaches	Beaches in the most popular tier, sampled one month before swim season.	watersgeo.epa.gov
Hospital quality	State's ranking in health care quality. State data is applied to each CZ in the state. Scale: 1–10, 10 being highest quality.	usnews.com
Annual precipitation (inches)	Annual inches of precipitation. Aggregated to CZ level from county level.	betsplaces.net
School expenditure per student	Average expenditures per student in public schools, USD per annum.	Opportunity Atlas

Table A.II: Empirical relationship between amenities and rent

Dependent variable: Annual rent in $	Coefficient estimate	Standard error	P-value
Median household wage (passthrough rate, ϵ)	0.204	0.006	0.000
Summer climate score	386.234	38.398	0.000
Winter climate score	1169.859	52.263	0.000
Population density (people/mile^2)	0.5519	0.520	0.000
Local government expenditures ($ per capita)	0.342	0.083	0.000
Crime rate (per mio. people)	-38.293	16.934	0.024
Walkability score	30.966	3.420	0.000
Beach access (total length in miles)	42.267	5.759	0.000
No. of top tier beaches	33.661	10.887	0.002
Hospital quality	98.093	19.276	0.000
Annual precipitation (inches)	-7.629	3.935	0.053
School expenditure per student ($ per annum)	0.126	0.052	0.015

Notes.– Displayed are coefficient estimates of the relationship between amenities and rents on the CZ level (equation (11)). Summer climate score and Winter climate score and Hospital quality are recorded on a 1–10 scale, 10 denoting the best possible score. The Walkability score is recorded on a 0–100 scale, 100 denoting highest walkability. The regression includes 690 commuting zones (out of a total of 740) for which all amenities and annual rents are observed in our data. The adjusted R^2 of the regression is 0.92.

geographically proximity and similarity in observables of CZs, we use the weighted euclidean norm

$$d_{CZ}(r, r') = \sqrt{\frac{1}{J} \sum_j \left(\frac{x_j(r) - x_j(r')}{\omega_j} \right)^2},$$

where $x_1(r)$ and $x_2(r)$ are AEA latitude and longitude and $\{x_j(r)\}_{j=3}^J$ are the remainder characteristics of CZ r. The weights $\{\omega_j\}$ correspond to the standard deviations of the observed characteristics across CZs, and across all pairwise combinations of CZs for the AEA coordinates.

Equipped with d_{CZ} and some cutoff $\bar{d}$, our algorithm proceeds as follows:

1. Start from the full list of CZs. Sort it in ascending order in terms of population size. Denote the resulting sorted list by $I = (r_0, r_1, ..., r_K)$. Start from $k = 0$.

2. Find $l^* = \underset{l>k}{\operatorname{argmin}}\, d_{CZ}(r_k, r_l)$.

3. If $d_{CZ}(r_k, r_{l*}) \leq \bar{d}$:

   (a) Merge r_{l^*} with r_k (removing r_{l^*} from I). Update the characteristics of the newly merged CZ, $\{x_j(r_k)\}_{j=1}^J$, by summing population sizes and type-$\boldsymbol{o}$ household population counts of r_k and r_{l^*}, and taking population weighted averages of all other characteristics.

(b) Iterate through $\{r_m : r_m \in I, m < k\}$ and merge each element r_m satisfying $d_{CZ}(r_m, r_k) \leq \bar{d}$ with r_k, following the steps outlined above to update the characteristics, $\{x_j(r_k)\}_{j=1}^{J}$, and removing r_m from I.

4. Increase k by 1 and proceed to the next element in I. Repeat steps 2–4 until $k = K$. The resulting I contains the list of merged CZs.

We set the cutoff to $\bar{d} = \frac{1}{3}$, implying that the most that two commuting zones can differ in any one characteristic is $\frac{1}{3}$ standard deviation. Starting from 690 CZs with non-missing data, out algorithm delivers a set of 517 merged CZs. The average geographic distance between CZ's centroids merged by our algorithm is 111 kilometers. The average distance in terms of d_{CZ} is 0.2.