How Local Are U.S. Labor Markets?: Using an Assignment Model to Forecast the Geographic and Skill Incidence of Local Labor Demand Shocks

Richard K. Mansfield*
University of Colorado-Boulder

March 22, 2019

Abstract

This paper examines how spatial frictions in labor markets differ by worker skill type and establishment industry, size, and skill requirements, and how such frictions shape the geographic and skill incidence of alternative local labor demand shocks, with implications for the appropriate level of government at which to fund “local” economic initiatives. LEHD data capturing the near universe of U.S. job transitions from 19 states facilitate the estimation of a rich two-sided assignment model of the labor market featuring thousands of parameters. The model is then used to generate simulated forecasts of many alternative local shocks featuring different establishment compositions. These forecasts suggest that existing local workers from the targeted public-use microdata area (encompassing at least 100,000 workers) account for only 7.0% (8.1%) of total welfare (employment) gains from stimulus shocks adding 500 jobs to a particular census tract, with at least 41.6% (34.5%) of welfare (employment) gains accruing to out-of-state workers. This is despite the fact that projected welfare and employment rate increases from a typical positive shock are 3 times larger for existing workers in the targeted Census tract than for workers from an adjacent tract, because workers in the target tract (or even the target PUMA) are a minuscule share of the national labor market. Further, the projected earnings incidence across local skill groups is quite sensitive to the shock’s establishment type composition, though alternative compositions produce increasingly similar incidence across skill groups at greater distances from the shock.

*This paper has benefitted greatly from an extraordinary effort by Henry Hyatt, as well as from considerable earlier assistance from Evan Buntrock, John Abowd, and Erika McEntarfer and valuable comments from Jeronimo Carballo, Terra McKinnish, Brian Cadena, Murat Iyigun, and participants at the NBER Trade and Geography Conference, the IZA World of Labor Conference, the Urban Economics Association, and multiple seminars at the University of Colorado-Boulder. This research uses data from the US Census Bureau’s Longitudinal Employer-Household Dynamics Program, which was partially supported by NSF grants SES-9978093, SES-0339191, and ITR-0427889; National Institute on Aging grant AG018854; and grants from the Alfred P. Sloan Foundation. All results have been reviewed to ensure that no confidential information is disclosed.
1 Introduction

Billions of dollars in local aid are spent each year by state, federal, and local agencies to support city-level or county-level economic development initiatives that seek to enhance labor market opportunities for workers in a particular skill class who live or work within the local jurisdiction (Bartik (2004)). These often take the form of local infrastructure spending, discounted loans or subsidies aimed at startup companies, or tax breaks to lure firms to relocate. In order to determine which types of firms or projects to support, federal, state, and local policymakers must predict not only which types of workers from which locations would be directly hired by the tax-supported firms, but also how the resulting ripple effects that operate through vacancy chains and pressure on local wages would indirectly benefit both local and more distant workers. In particular, whether to fund such initiatives at the city, county, or state level depends critically on the shares of the initiative’s welfare incidence expected to redound to workers within the city, county, and state borders, respectively. Which types of targeted firms will yield a geographically concentrated impact in which the labor demand shock trickles down to lower skill levels rather than out to more distant locations?

While a large literature in economics seeks to evaluate the incidence of place-based labor demand policies and shocks, most reduced-form methods focus on quite local impacts, with more distant towns, counties or states either excluded from the sample or used as control groups, thereby ignoring the possibility that these more distant areas might collectively account for a sizeable share of shock incidence, even if no one particular area is strongly affected. Furthermore, by virtue of their focus on particular policies or shocks occurring in one or a small number of locations, these studies are usually ill-equipped to compare the incidence of shocks featuring different demand compositions or to examine differential skill incidence among local and less local areas (due to small samples of workers within a small radius around the shock and/or a lack of detailed data on distant locations).

The primary difficulty is that either evaluating or predicting worker-level welfare incidence across a variety of alternative local labor market shocks requires a spatial equilibrium model that accommodates ripple effects by incorporating the network of spatial linkages among overlapping local labor markets while simultaneously featuring heterogeneity in worker and firm preferences, search costs, and match productivities along a variety of observable dimensions.

Motivated by this challenge, this paper makes two central contributions. First, I develop a theoretically-motivated empirical framework for assessing and forecasting welfare incidence across location-by-demographic group categories from labor demand shocks featuring alternative geographic and establishment type compositions (or worker type compositions for labor supply shocks). I do this by adapting to the local labor market setting the two-sided assignment game analyzed originally by Koopmans and Beckmann (1957) and
Shapley and Shubik (1972) and whose empirical implications were highlighted in the marriage market context by Choo and Siow (2006). Second, after estimating the parameters of the model, I analyze a large set of model simulations that illustrate several general properties of local labor markets in the United States that effectively create a useful national prior about which types of workers are most sensitive to which types of local labor demand shocks.

Several key features of Choo and Siow (2006)’s version of the assignment game facilitate these goals. First, it can accommodate multidimensional heterogeneity based on unordered categorical characteristics for agents on both sides of the matching market. In particular, this allows the model to accommodate arbitrary spatial links between workers and establishments in different geographic units, including geographic units of both very small and large sizes. It also permits analysis of incidence across demographic groups such as races, age groups, or industries (or combinations thereof) without requiring any hierarchical ordering.

Second, the assignment game requires market clearing, optimizing behavior by all market agents, and explicit payoffs to each agent from each possible job match, making it well-suited for forecasting welfare effects from exogenous shocks. Third, the key parameters of the model (mean relative joint surpluses among matched pairs of workers and firms belonging to observable types) can be identified from a single cross-sectional labor-market transition between origin and destination allocations of workers to jobs, and are sufficient to perform counterfactuals that yield the resulting allocation and impact on payoffs for all players (workers and firms) from any arbitrary change in the composition of labor supply, labor demand, or both.

Finally, these counterfactuals do not require the specification of a more fundamental structural model of utility, firm production, and moving costs, ensuring that none of the heterogeneity present in the transition patterns is lost in paring down to a small number of interpretable structural parameters. The downside to such a “sufficient statistic” approach is that the set of counterfactuals that can be performed is limited to those involving exogenous changes in either the type composition of labor supply and/or demand or composite joint surplus parameters. Furthermore, while the heterogeneity on both sides of the labor market can be modeled much more richly than in other structural models, the housing and product markets are not explicitly modeled (though their impact may nonetheless be captured by the estimated surplus parameters through the way they affect job-to-job flows). Thus, the estimated welfare changes only capture “labor-related” welfare changes, and should be thought of as complementary inputs to local policy decisions along with estimates of house price and product price elasticities. For example, local residents and policymakers who are concerned that local development initiatives creating new high-skilled positions might lead to potential rent increases for low-skilled renters might wish to know whether downstream earnings and employment opportunities for such renters will increase enough to compensate.
I estimate the model and perform a variety of counterfactual simulations using matched employer-employee data from the Longitudinal Employer-Household Dynamics (LEHD) database on a subset of 19 U.S. states that approved the use of their employment records. The data display three key properties that make it suitable for these forecasts. Namely, 1) they capture the (near) universe of job matches from the participating states, mitigating selection problems, 2) they include hundreds of millions of job matches, allowing precise estimates of the large number of parameters necessary to capture the complex two-sided multidimensional sorting that occurs in the labor market, and 3) workers’ establishments are geocoded to the census tract level. These properties, when combined, make it feasible to study incidence across worker types at the very local level necessary to make the estimates useful to local policymakers, while still allowing for complex spatial ties between the local area and the surrounding towns, counties, and states that make the estimates useful to state and federal policymakers. In particular, these data, when combined with the assignment model, provide the necessary inputs for computing the shares of employment and welfare gains or losses from alternative local labor demand shocks that accrue to workers of particular skill levels located within particular radii around the shock.

The counterfactual simulations involve establishment relocations or stimulus projects that create new job positions in particular U.S. locations (census tracts) featuring alternative combinations of establishment size, average pay, and industry supersector. I also consider “natural disaster” simulations akin to a tornado or flood that eliminate a share of all jobs in a particular geographic location.

I find that welfare gains or losses from very local shocks are widely distributed. For example, the simulations suggest that as little as 7.0% of the job-related utility gains and 8.1% of the net employment gains from a stimulus project that creates 500 new jobs in a given census tract accrue to workers already working (or seeking a job) in the surrounding PUMA\(^1\) at the beginning of the year, while at least 41.6% (34.5%) of the job-related utility (employment) gains accrue to workers beginning the year outside the state.

Such geographic dispersion of welfare gains occurs despite the fact that the same simulations suggest that a randomly chosen worker in the targeted census tract is about 3, 27 and over 9,000 times more likely to fill one of the new vacancies than a randomly chosen worker in an adjacent tract, in an adjacent PUMA, and in a non-adjacent state, respectively. These seemingly inconsistent findings are the result of two mechanisms.

First, since many of the workers likely to join the incoming firms were already employed, so that their transitions generate further openings for others, the share of the stimulus jobs taken by more vs. less local workers dramatically overstates the local concentration of the overall employment and welfare incidence of the labor demand shock. Second (and more

\(^1\)PUMAs or “public-use microdata areas” are mutually exclusive and exhaustive collections of contiguous counties and census tracts encompassing at least 100,000 residents. They are used in this paper as a geographic unit that captures a small city-sized population regardless of nearby population density.
importantly), because a single census tract generally only features a few thousand workers, its workforce makes up a very small share of the national labor market: there were 73,057 tracts and 2,378 PUMAs defined in the 2010 census. Consequently, even quite substantial and disproportionate welfare gains for the most local workers cannot account for more than a tiny share of the aggregate welfare gains. Indeed, the predicted utility and employment gains for an initially local worker are 2.6 (3.2), 3.8 (8), and 12.3 (448) times as large as for a randomly chosen worker in an adjacent tract, in an adjacent PUMA, and in a non-adjacent state, respectively.

Workers initially working in the focal tract receive an estimated $1,045 increase (in 2011 dollars) in money metric utility from the typical stimulus package (relative to the least affected location/skill category combination), while workers initially working 1, 2, and 3 or more tracts away receive expected utility gains of $395, $278, and $164 respectively. Workers initially working 1, 2, and 3+ PUMAs away within the state receive the utility equivalent of $164, $143, and $109 in annual earnings gains, while workers one and 2+ states from the site of the shock receive average gains of $89 and $85.

Averaging across simulations, the results suggest that among the most local workers, the utility gains are largest among the initially high-paid workers ($1242 in money-metric utility units) and prime-age non-employed ($1165), and smallest among initially low-paid workers ($999) and the young non-employed ($620), where I use initial earnings as a proxy for worker skill. However, these averages mask substantial heterogeneity in projected impacts across shocks featuring different establishment compositions and across sites of simulated shocks. The average money metric utility gains for the same four skill/age groups are $1728, $1015, $825 and $415 for stimuli consisting of jobs at small, previously high paying establishments vs. $872, $1297, $1130 and $802 for stimuli featuring large, previously low paying establishments. I also find that demand shocks consisting of additional jobs at small, low-paying establishments in the other services supersector generate the most locally concentrated employment impact for initially low-paid local workers ($1703), while stimuli featuring jobs at large, high-paying information supersector establishments generate the smallest local employment impact ($718).

Interestingly, regardless of the establishment composition, as the simulated shocks ripple outward, they becomes less and less skill-biased: predicted differences in welfare (or employment) gains among skill groups converge as one consider workers at initial locations further from the site of the shock. I also find that the share of employment gains from stimuli that accrue to workers initially working or seeking a job within the chosen PUMA is twice as high in rural areas as in urban areas (≈15% vs. 7.5%), and that requiring the newly created jobs to be filled exclusively from existing PUMA workers (or jobseekers) dramatically increases the share of employment gains that accrue to such workers (to nearly 25% from less than 10%).

To accommodate such multidimensional heterogeneity, the parameters that govern the
impact of simulated shocks exploit for identification the full sample of worker flows rather than exclusively flows involving locations experiencing local employment shocks. To show that the model is nonetheless capable of generating accurate forecasts for moderately sized local shocks, I also perform a model validation exercise in which parameters estimated using data from the previous pair of years are used to predict the reallocation that actually occurred among 514 census tracts that experienced gains or losses of between 100 and 3000 jobs within one year between 1996 and 2010. When job transitions are assigned to one of 10,752 categories defined by worker initial earnings and the distances of the origin and destination locations from the shock, a dissimilarity index reveals that on average only 0.9% of national job transitions and 5.3% of transitions originating in the targeted PUMA would need to be assigned to a different category in the simulation to perfectly match the post-shock reallocation that actually occurred.

This paper builds primarily on three literatures. The first consists of evaluations of particular place-based policies or local economic shocks. Most papers in this branch use average wages or employment rates in the targeted location as the outcome of interest, seek to define a control group of alternative locations, and evaluate the policy or shock’s impact using a treatment effect framework. This literature is vast, and is thoroughly discussed by survey articles such as Glaeser et al. (2008), Moretti (2010), Kline and Moretti (2013), and Neumark and Simpson (2014). A pair of papers in particular stand out.

Autor et al. (2014)’s evaluation of the worker-level impact of China’s accession to the WTO is notable for its attention to heterogeneity in incidence across demographic and skill groups. They find that the negative import competition shock particularly affected the cumulative earnings of those with low initial earnings or limited labor force attachment. However, because they consider local variation in the incidence of a national-level shock, their estimates do not provide much guidance on the geographic incidence of a small but geographically concentrated demand shock.

Busso et al. (2013)’s evaluation of the U.S. empowerment zone system also stands out as one of the few quasi-experimental papers to explicitly evaluate social welfare impact, which they accomplish by deriving a set of sufficient elasticity parameters that can be cleanly identified. Interestingly, they find that while empowerment zones significantly increase wages and employment of zone residents, they do not meaningfully affect rent prices. This suggests that for very local shocks where commuting adjustments play a key role in facilitating the shock response, the impact on rent need not be first-order.

---

2A particularly prominent paper in this branch of the literature is Greenstone et al. (2010), who compare employment gains in counties making winning bids for “million-dollar” plants to control counties who made losing bids. More recent contributions include Gregory (2013), Freedman (2013), and LeGower and Walsh (2017).

3The authors point out that the limited rent price impact might be due to the particularly depressed nature of the targeted locations, which could make them undesirable residential locations (or subject to rent control).
The paper also contributes to a fast-growing literature on structural spatial equilibrium models designed to forecast the incidence of economic shocks across spatially-linked geographic areas. Several recent papers in particular deserve discussion.

Caliendo et al. (2015) (hereafter CDP) consider the geographic and sectoral incidence of China WTO entry. They develop a full dynamic general equilibrium framework that incorporates input-output linkages in goods markets as well as labor market linkages among a system of 50 U.S. states and 37 countries. They show how counterfactual dynamic equilibrium paths can be evaluated for alternative structural shocks (changes in trade costs, mobility costs, productivities) without estimating all the primitives of the model. The present paper relies on a very similar “sufficient statistics” approach, in that it evaluates the distribution of welfare impacts from demand shocks of alternative compositions without identifying the fundamental utility, production function, and moving cost parameters of the structural model. The model below imposes even less structure on the form of production and utility than CDP, but is also more limited in the set of counterfactuals it can evaluate.

Monte et al. (2015) highlight the roles of commuting vs. residential mobility in clearing U.S. labor markets across geographic space and in determining the incidence of local labor demand shocks. Like CDP, they use a trade-theoretic approach to model the joint choice of residential and work location among U.S. counties, and incorporate commuting costs, local amenities, and geographic trade costs. They show that a richer structural model that incorporates the network of commuting flows can better predict the heterogeneity in incidence of the million dollar plant openings evaluated by Greenstone et al. (2010).

Schmutz and Sidibe (2016) adapt a search and matching framework in the style of McCall (1970) to model the importance of spatial linkages in determining the incidence of local shocks. Exploiting worker flows among a system of French metropolitan areas, they find that search frictions play a greater role than moving costs in limiting worker mobility, suggesting the potential promise of efforts to disseminate information about distant jobs.

Each of these papers aggregates locations to at least the county level, leaving considerable room for an analysis of the geographic incidence of very local shocks of the type considered by policymakers in particular towns or cities. Manning and Petrangelo (2017), by contrast, represents the most notable attempt to determine the equilibrium incidence across nearby areas of small scale shocks. Like Schmutz and Sidibe (2016), they propose a search and matching model and fit the model-predicted geographic distribution of vacancy outflows to data on changes in vacancy stocks from local job search centers in Britain. Like this paper, they simulate the impact on the geographic distribution of unemployment of an exogenous increase in vacancies (new jobs) within particular census wards (similar in size to the census tracts used here). They also find evidence that labor markets are quite local, in the sense that moderate distance to vacancies substantially decreases the probability of

---

4Due to a lack of micro-level residential data, the present paper does not consider whether the new jobs formed by job-to-job transitions involve residential mobility.
an application. Nonetheless, they also find that ripple effects from overlapping markets cause the unemployment incidence to spread widely, with very little of the employment gain accruing to the ward receiving the shock (less than is reported here).

While the present paper lacks the explicit housing and product markets modeled in CDP, the commuting links modeled in Monte et al. (2015), and the distinction between search and moving costs highlighted by Schmutz and Sidibe (2016) and Manning and Petrongolo (2017), it features a much richer labor market. Among the papers just described, none feature any heterogeneity in worker productivity (or any observable worker heterogeneity beyond initial location), and only CDP (industry differences) features any observable firm heterogeneity besides firm location. Because none of these models feature multidimensional two-sided sorting, the model featured in this paper is the only one equipped to evaluate differential incidence both across space and across skill/demographic groups from local labor demand shocks with alternative firm compositions.⁵

Finally, this paper also builds upon and draws heavily from the theoretical literature on the identification and estimation of two-sided assignment games. To my knowledge this is the first large-scale labor market application of a two-sided assignment model.⁶ The theoretical properties of such assignment games have been well-established for at least a generation.⁷ However, the empirical content of the model for contexts in which the universe (or a large random sample) of all market entrants on both sides and their matches can be observed has only recently attracted interest, with Choo and Siow (2006)’s pioneering paper leading to contributions by Chiappori and Salanié (2016), Menzel (2015), and Galichon and Salanié (2015), among others. I make two contributions to this theoretical literature.

First, I consider implementation in a context with a very large number of match observations and a very large number of types on both the supply and demand side. I address the challenge of a somewhat sparse matching matrix by introducing a smoothing procedure designed to aggregate matching patterns across “nearby” match types without smoothing away the heterogeneity the model is designed to highlight.

Second, I consider the limits to identification in a context where the number of unmatched partners of each type is either unobserved or only observed on one side of the market: while nonemployment may be inferred with reasonable accuracy in the LEHD data, unfilled vacancies are absent. The existing identification strategies employed by Choo and Siow (2006) and Menzel (2015), among others, rely heavily on observing the number of singles on both sides. I discuss conditions under which ignoring unmatched partners would

---

⁵The CDP and Monte et al. (2015) frameworks could accommodate such features, but at the cost of considerable computational burden and model complexity.

⁶See Tervio (2008) and Chen (2017) for applications of the assignment game to the market for CEOs. See also Lindenlaub (2017) for an application of the continuous (hedonic) analogue to the discrete assignment game to the matching of workers and occupations.

⁷See Koopmans and Beckmann (1957), Shapley and Shubik (1972), Roth and Sotomayor (1992), and Sattinger (1993) among others.
not affect the incidence of policy interventions among originally matched agents.

The rest of the paper proceeds as follows. Section 2 describes the two-sided assignment game that forms the theoretical basis for the empirical analysis. Section 3 illustrates how to apply the insights of Choo and Siow (2006) to the context of labor market transitions to identify a set of joint surplus parameters that are sufficient to perform counterfactual simulations of labor demand shock incidence. Section 4 describes the LEHD database and presents summary statistics that motivate the subsequent analysis. Section 5 describes sample selection and the smoothing procedure used to eliminate disclosure risk and minimize the sparsity of the large-scale transition matrix whose entries determine the surplus parameters. In addition, Section 5 also provides detail about the particular specifications of labor demand shocks that I simulate, the procedure used to perform the simulations, and the methods used to aggregate the resulting counterfactual allocations of workers to positions into interpretable statistics that effectively characterize variation in shock incidence. Section 6 presents the main findings, and Section 7 concludes the paper.

2 Model

2.1 Model Overview

In this section I model the evolution of the labor market across adjacent time periods as a static cooperative matching game played by workers and firms. The model is based on Choo and Siow (2006)’s model of the marriage matching market, but introduces a number of features and extensions necessary to adapt their model to a labor market setting. The exposition of the model closely mirrors Galichon and Salanié (2015), which generalizes Choo and Siow (2006). Section 2.2 lays out the basics of the matching game. Section 2.3 describes how the workers and positions (the game’s agents) and the job matches that determine the game’s payoffs are aggregated to types and groups, respectively. Section 3 imposes additional structure on the model that facilitates the identification and estimation of the underlying group-level match surpluses that determine the frequencies of particular kinds of job transitions. Section 3.2 shows how these estimated match surpluses can be used to construct counterfactual simulations capturing the incidence of local labor supply and demand shocks of varying worker and establishment compositions.

2.2 Defining the Assignment Game

Suppose that in a given year $y$ there are $I$ potential workers participating in the labor market, with the set of individual workers denoted $\mathcal{I}$. Each worker begins the year in a job match with a position $j(i)$ at establishment $m(j(i))$, determined in year $y - 1$, from the set of possible positions $\mathcal{J}$. Let positions featuring $m(j) = 0$ represent unemployment so that positing an initial “job” match for each worker is without loss of generality.
The value to worker $i$, currently at position $j$, of accepting a position $k$ the following year is denoted $U(i,j,k)$. The worker’s potential annual earnings in year $y$ if they accept position $k$, denoted $w_{ik}$, is assumed to be additively separable from all other determinants of the worker’s payoff, so that $U_{i,j,k}$ takes on a quasi-linear money-metric form: \[^8\]

$$U(i,j,k) = \pi_{ijk} + w_{ik}$$

(1)

$\pi_{ijk}$ captures the combined value to worker $i$ of a variety of payoff components. I show below that the researcher need not specify any of the fundamental parameters or the functions governing their links to payoffs in order to construct the counterfactual simulations that form the primary contribution of the paper. Any value function specifications in which current worker earnings are additively separable from other payoff determinants will suffice. That said, careful thought about which determinants of the payoff are likely to be large in magnitude and differential across alternative workers, positions, and job matches will be important in guiding the choice of characteristics used to aggregate workers and positions to types in section 2.3 below, as well as for evaluating the plausibility of assumptions underlying the counterfactual simulations that are laid out in section 3.2 below.

Such components might include worker $i$’s particular valuation of various non-pecuniary amenities offered by position $k$ (including the desirability of the geographic location of the position), as well as any search, moving, or training costs paid by worker $i$ associated with finding, moving to, and settling into position $k$ from position $j$. \[^9\] They might also include the continuation value stemming from the fact that worker $i$ will begin the following year as an incumbent, trained worker at position $k$, which might depend on the productivity gains from firm-specific experience and the availability of other opportunities in position $k$’s local labor market. From this point forward, because $j$ is fully determined by $i$ in the static assignment game I consider, I will frequently suppress $j$ and replace $U(i,j,k)$ with $U(i,k)$, and will do the same with other functions that depend on $j$, except where necessary to illustrate the role of workers’ origin job assignments in determining the reallocation of workers across job matches.

On the other side of the market there are $K$ potential positions at establishments that seek workers in year $y$ that make up the set $\mathcal{K}$. Note that the intersection of the sets $\mathcal{K}$ and $\mathcal{J}$ may be quite large, so that many of the end-of-period positions in $\mathcal{K}$ can potentially

---

\[^8\]Since I have data on annual earnings but not wages or hours, for simplicity I assume that the hours associated with a job match are fixed by contract and common across positions for a given worker, and focus exclusively on earnings.

\[^9\]The traditional assignment game does not feature any stochastic search frictions, so that each agent might in principle match with any agent on the opposite side of the market. However, Menzel (2015) shows that one can introduce a probability $r(i,k)$ that $i$ and $k$ meet that is independent of other payoff determinants, assign the joint surplus from the match to $-\infty$ if the pair does not meet, and use these alternative payoffs to determine the stable matching. Alternatively, search costs might be modeled as a deterministic cost that must be paid by an agent to an intermediary (e.g. a headhunter or a job matching website) to reveal/allow contact with particular agents on the opposite side of the market (without which they cannot be found).
be “filled” by simply continuing a job match that already exists. I assume that each establishment makes independent hiring decisions for each position, so that I can model the preferences of positions over individual workers rather than modeling establishment preferences over collections of workers.\footnote{One justification for treating positions as independent is that there are nontrivial costs of coordinating multiple independent hires/retention decisions that outweigh the gains from better exploiting the complementarities that exist in the production process. See Roth and Sotomayor (1992) for a detailed analysis of how the properties of the assignment game change when establishments have preferences over collections of workers.} Let the value of hiring (or retaining) a given worker $i$ for a particular position $k$ in establishment $m(k)$ be given by $V(i,k)$. The potential annual earnings paid by $k$ to worker $i$ in year $t$ is again assumed to be additively separable from all other determinants of the position’s payoff, so that $V(i,k)$ can be written as:

$$V(i,k) = \pi_{ik}^k - w_{ik} \quad (2)$$

Akin to $\pi_{ik}^i$, $\pi_{ik}^k$ is likely to reflect the contributions of several payoff components, whose relationship need not be fully specified by a particular function form. These components might include the contribution of worker $i$ to establishment $m(k)$’s revenue in the coming year, any recruiting, moving, and training costs borne by establishment $m(k)$ in hiring worker $i$, as well as any discounted continuation value to establishment $m(k)$ from beginning the next year with worker $i$ already installed in position $k$, including the fact that retaining $i$ next year would not require further recruiting/training costs.

One can then define the joint surplus from the transition of worker $i$ to position $k$ as the sum of the worker and position valuations of the transition:

$$\pi_{ik} \equiv U(i,k) + V(i,k) = \pi_{ik}^i + \pi_{ik}^k \quad (3)$$

Since worker annual earnings in the current period are additively separable in both the worker’s and position’s payoffs, this assignment model exhibits transferable utility, with current annual earnings representing the transfer. Written in this form, one can see that the game has the exact structure of the classic assignment game analyzed by Koopmans and Beckmann (1957) and Shapley and Shubik (1972).

A matching or market-wide transition in this labor market is an $I \times K$ transition matrix $\mu$ such that $\mu_{i,k} = 1$ if worker $i$ matches with position $k$ at the end of the period, and 0 otherwise. As in Galichon and Salanié (2015), I focus on stable matchings, which have the property that a division of joint surplus exists in each proposed job match such that no currently unmatched worker-position pair can find any division of the joint surplus from their potential match that makes both the worker and position strictly better off than they are under the proposed matching. Shapley and Shubik (1972) showed that the set of stable matchings coincides with both the core of the assignment game and with the set of competitive equilibria from a decentralized labor market. Furthermore, they show that in
the presence of transferable utility there will exist a unique assignment (or, equivalently, competitive equilibrium allocation) of origin job matches to destination job matches as long as preferences are strict on both sides of the market. This equilibrium allocation/stable assignment will maximize the aggregate surplus, and thus can be found by solving a linear programming problem.\footnote{The joint surplus is given by \( \sum_{i \in I} \sum_{k \in K} \mu_{i,k} \pi_{i,k} + \sum_{i \in I} \sum_{k \in K} \mu_{i,k} \pi_{i,k} = 0 \) \( \forall \ k \mu_{i,0} \pi_{i,0} + \sum_{k \in K} \mu_{0,k} \pi_{0,k} \). Here, \( \mu_{i,0} \) and \( \pi_{i,0} \) denote worker \( i \)'s payoff from remaining unemployed and position \( k \)'s payoff from remaining vacant. The constraints stem from the fact that each position and each worker can be matched to at most one counterpart: \( \sum_{i' \in \mathcal{I}} \mu_{i',k} \leq 1 \forall \ k \in K \) and \( \sum_{i' \in \mathcal{I}} \mu_{i,k'} \leq 1 \forall \ i \in \mathcal{I} \).}

Equivalently, the unique stable assignment can also be found by solving the dual problem: identifying a set of worker discounted utility values \( \{r_i\} \) and position discounted profit values \( \{q_k\} \) that minimize the total “cost” of all workers and positions, \( \sum_{i \in I} r_i + \sum_{k \in K} q_k \), subject to the constraint that these values cannot violate the underlying joint surplus values: \( r_i + q_k \geq \pi_{i,k} \forall \ (i,k) \). This dual problem yields the following conditions that define the optimal assignment (Galichon and Salanié (2015)):

\[
\mu_{i,k} = 1 \text{ iff } k \in \arg \max_{k \in K \cup \{0\}} \pi_{i,k} - q_k \text{ and } i \in \arg \max_{i \in I \cup \{0\}} \pi_{i,k} - r_i \tag{4}
\]

I aggregate these conditions in the next section to deliver identification of aggregate “type”-level surpluses.

Finally, given optimal (or equilibrium) worker and position payoffs \( \{r_i\} \) and \( \{q_k\} \) from the dual solution, Shapley and Shubik (1972) show how to decentralize this optimal assignment via a set of earnings transfers \( w_{ik} \) that appear in the worker and position value functions above:

\[
w_{ik} = \pi_{i,k}^k - q_k \tag{5}
\]

Because \( r_i + q_k = \pi_{i,k} = \pi_{i,k}^i + \pi_{i,k}^k \) for any pair \( (i,k) \) that is matched in the unique stable match, this also implies that:

\[
w_{ik} = r_i - \pi_{i,k}^i \tag{6}
\]

Then the conditions (4) can be rewritten as the standard requirements that worker and establishment choices must be utility- and profit-maximizing, respectively:

\[
\mu_{i,k} = 1 \text{ iff } k \in \arg \max_{k \in K \cup \{0\}} \pi_{i,k}^k + w_{ik} \text{ and } i \in \arg \max_{i \in I \cup \{0\}} \pi_{i,k}^k - w_{ik} \tag{7}
\]

This shows that the market-clearing earnings amounts will in general be specific to worker-position pairs \( (i,k) \). By contrast, the market-clearing utilities \( r_i \) and profit contributions \( q_k \) will be worker-specific and position-specific, respectively, which is a property that will be exploited below. Importantly, while the stable assignment \( \mu_{i,k} \) is generally unique, the equilibrium payoffs and wages are not: all \( \{r_i\} \) values can generally be shifted either up or down by a small increment \( \delta \) (with offsetting decreases or increases for \( \{q_k\} \)) without
violating any of the stability conditions. The exact equilibrium payoffs/wages that emerge depend on the particular process by which the decentralized labor market converges.

While the analysis does not require taking a stand on a particular earnings determination process, it is nonetheless illuminating to present one candidate process: a simultaneous ascending auction. In such an auction, positions are bidding on all workers simultaneously. Each position may only be the highest bidder for a single worker (or for none, if it chooses to remain vacant). Workers may set reservation utilities which will vary based on the value different workers place on remaining unemployed for a year. The position \( k \) that bids the highest discounted utility \( r_i \) to a worker \( i \) retains the worker and pays the worker an annual earnings amount \( w_{ik} \) that, when combined with the non-pecuniary component \( \pi_{ik} \), is sufficient to ensure the worker’s promised valuation \( U_{ik} = r_i \). The auction ends when no position wishes to outbid any other position for a worker. Some workers may remain unemployed, and some positions may remain unfilled.

Importantly, under the quasi-linear utility specification above, even though positions are bidding values of a one period commitment \( U_{ik} \) (which include the value of beginning the next period as an incumbent at the position), they each start at different baseline levels of \( \pi_{ik} \), and changes in bidding can always take the form of annual earnings increases. Thus, when solving for changes in the stable assignment of workers to firms following shocks to labor demand composition below, one may compute the changes in reservation utilities \( r_i \) that clear the market, and one may scale these changes in terms of annual earnings gains (though in some cases workers achieving utility gains will prefer to take an earnings cut to work at an establishment offering non-pecuniary values that more than offset the earnings decrease).

2.3 Modeling the Match Surpluses

Part of the joint transition surplus \( \pi_{ik} \) for worker-position pair \((i, k)\) from transition \((i, j(i))\) to \((i, k)\) is likely to be common to any transition \((i', j(i'))\) to \((i', k')\) that shares certain salient characteristics of the worker, positions, origin or destination job matches, or even transition. For example, positions at larger firms may face smaller per-position costs of recruiting distant workers due to economies of scale; highly skilled workers may generate larger surplus at positions whose output is particularly sensitive to worker skill. Thus, I assign each potential transition \((i, k)\) to one of a set of mutually exclusive groups \( g \in G \) (with \( G \equiv |G| \)), and use the notation \( g(i, k) \) to denote the group to which transition \((i, k)\) has been assigned. Importantly, these groups are always defined by a combination of observable characteristics of the worker \( i \), the origin job \( j(i) \) or job match \((i, j(i))\), the destination position \( k \), or even the transition \((i, k)\). The characteristics that define the set of transition

\[ \text{12} \] Thus, when generating counterfactual allocations in the empirical work below, I will often normalize changes in utility values for one type of worker to be zero, and analyze relative changes that are better identified.
groups should be chosen to capture as comprehensively but parsimoniously as possible the underlying (structural) preferences, productivities, moving costs, and geographic search costs that enter into the joint surplus $\pi_{ik}$ and thus determine the relative desirability of the job match for both the worker and the position.

Some subset of these observed characteristics may only relate to the worker $i$, the origin position $j(i)$, or the worker-origin position job match $(i, j(i))$, and will be common to all destination positions $k' \in K$. I use this subset of characteristics to assign each worker-origin position job match $(i, j(i))$ to an origin type $o \in \mathcal{O}$, and use the notation $o(i)$ or $o(i, j(i))$ to denote the origin type to which $(i, j(i))$ has been assigned. In the empirical work below, the origin types are defined by unique combinations of a pair of characteristics. The first is the geographic location (either census tract, public-use micro area PUMA, or U.S. state) of the establishment $j$ at which worker $i$ works in the origin period (or most recently worked for initially unemployed workers). The second is the skill level of worker $i$, which is proxied by the national earnings quartile associated with $i$’s earnings while at position $j$ (in year $y-1$), if employed in $y-1$. If the worker was unemployed in period $y-1$, I assign them to one of two unemployment categories, based on age ($< 25$ vs. $\geq 25$), to distinguish relatively new entrants or recent graduates from experienced unemployed workers.\textsuperscript{13}

Analogously, another (mutually exclusive) subset of the observed characteristics defining the transition group $g$ may only characterize the destination position (i.e. they are common to all origin workers $i$). This subset is used to assign each destination position $k' \in K$ to a destination type $d \in \mathcal{D}$, denoted $d(k)$. In the empirical work, the destination types are defined by unique combinations of the following characteristics: the geographic location (U.S. census tract, PUMA, or state) of establishment $m(k)$, the quartile to which establishment $m(k)$’s total employment belongs (in the origin year) in the national establishment-level employment distribution, the quartile to which establishment $m(k)$ belongs (in the origin year) in the national establishment-level average earnings distribution (intended to proxy for the average required skill level of establishment $m(k)$), and the industry supersector of establishment $m(k)$.

Finally, let $z(i, j(i), k) \equiv z(i, k)$ capture the remaining subset of characteristics defining the transition group that depend on both $(i, j(i))$ and $k$. In the empirical work below, the

\textsuperscript{13}Since the goal is to characterize the geographic scope of workers’ searches for jobs and firms’ searches for employees, I would ideally use residential location to define the origin job type and establishment location to define the destination job type. However, I do not have access to the residential location of the worker, so I use establishment location in year $y-1$ as a proxy for residential location in year $y-1$.

\textsuperscript{14}Note that I interpret the earnings category of $i$ in year $y-1$ as a proxy for worker $i$’s skill, but am interpreting the average earnings category of establishment $m(k)$ in year $y$ as a proxy for the skill requirements of position $k$. This seeming inconsistency can be rationalized if one assumes that the worker, once hired, develops the skills required by the job by the end of the year (perhaps incurring some training costs that could be borne by either the worker or the establishment and thus affect the joint surplus $\pi_{ik}$). The assignment model also suggests that past equilibrium transfers will generally reflect both the skill requirements of the position and the skill requirements of the worker, so that they may be effective proxies for both in the absence of other information.
single $z$ characteristic will be an indicator for whether the “transition” represents continued employment at the same establishment, $1(m(k) = m(j(i)))$, and is intended to capture the fact that search, recruiting, and training costs do not have to be repaid by existing workers. This allows the placement of job stayers and job movers into different groups, which in turn allows establishments to retain existing employees at different rates than they hire other local workers (important for predicting which workers ultimately accept newly created jobs).\footnote{Mourifié and Siow (2017) use the same approach to distinguish marriage from cohabitation among relationships featuring particular male and female types in their marriage market setting.}

Thus, without loss of generality one can rewrite the mapping $g(i, k)$ as $g(o(i), d(k), z(i, k)) \equiv g(o, d, z)$. Importantly, while knowledge of $g$ is sufficient to recover both $o$ and $d$, knowledge of $o$ and $d$ need not uniquely identify the group $g$ (due to the presence of $z$). In a slight abuse of notation, I will sometimes use $o(g) = o(g(i, k)) = o(i)$ to refer to the origin type associated with group $g$, and I will use $d(g) = d(g(i, k)) = d(k)$ to refer to the destination type associated with group $g$.

Given these definitions, one can decompose the transition surplus $\pi_{ik}$ into the part that is common to all transitions classified as group $g(i, k)$, denoted $\theta_g$, and an idiosyncratic component $\epsilon_{ik}$ specific to the particular transition $(i, k)$:

$$\pi_{ik} = \theta_{g(i,k)} + \sigma \epsilon_{ik} \quad (8)$$

$\epsilon_{ik}$ might reflect, for example, the low psychic costs of a particular worker who is moving back to the location where his family lives, or perhaps particular skill requirements of position $k$ that worker $i$ uniquely possesses. Following Decker et al. (2013), $\sigma$ is a scaling parameter that captures the relative importance of idiosyncratic components of the matching surplus compared to components that are common among all transitions classified into the same group $g$ in determining the variation in match surpluses across potential pairs $(i, k) \in I \times K$. As will be shown below, the introduction of $\sigma$ will not change the unique stable job assignments in the counterfactual simulations, but it will play a key role in determining the size of changes in offered utility values $r_i$ for particular workers in particular locations that are necessary to facilitate the reallocation that yields the stable assignment.

The goal is to use the observed matching $\mu$ to recover the set of group mean surplus values $\{\theta_g\}$. As Galichon and Salanié (2015) emphasize, one could impose further structure on the production, utility, search cost, and recruiting cost functions that comprise the joint surpluses, so that $\theta_g$ could be written as $\theta_g(\lambda_1)$ for some smaller set of structural parameters $\lambda_1$, with the distribution of $\epsilon_{ik}$ depending on a second parameter set $\lambda_2$. Maximum likelihood could then be used to relate the observed match $\mu$ to the parameters of the model.

Driven by a combination of computational considerations and an interest in being agnostic about the various structural functions that underlie $\{\theta_g\}$, I follow Choo and Siow
(2006) and leave the set \(\{\theta_g\}\) unrestricted, achieving identification instead by assuming that \(\epsilon_{ik}\) is independent and identically distributed across all alternative matches \((i, k')\) and \((i', k)\) \(\in I \times K\) and follows a Type 1 extreme value distribution. Unlike Choo and Siow (2006) and Galichon and Salanié (2015), I allow for the possibility that part of the match surplus is truly idiosyncratic: the combined surplus from two transitions \((i k)\) and \((i'k')\) would be altered if the two workers swapped destination positions, even if \(i\) and \(i'\) are both associated with the same origin type \((o(i) = o(i'))\) and \(k\) and \(k'\) are both associated with the same destination type \((d(k) = d(k'))\). Given the coarseness of the origin and destination types in the empirical work, such within-type-combination heterogeneity in match quality is very likely to exist in the labor market. Allowing such match-level heterogeneity, however, comes at a cost: as discussed in Section 3.2 and Appendix A4, I forfeit a straightforward way to use observed transfers to separate the group mean surplus \(\theta_g\) into group-level worker and position subcomponents \(\theta^l_g\) and \(\theta^f_g\) analogous to the transition-level surplus components \(\pi_{ik}\) and \(\pi_{ik}'\) defined above. Fortunately, this further decomposition is not necessary to perform an important class of counterfactual simulations (discussed further below).

3 Identification

3.1 Identification of the Set of Group-Level Match Surpluses \(\{\theta_g\}\)

Recall from section 2.2 that a necessary condition for a matching \(\mu\) to be stable (and thus sustainable as a competitive equilibrium) is that there exists a set of worker values \(\{r_i\}\) such that \(\mu_{ik} = 1\) implies that \(i \in \arg \max_{i \in I} \pi_{ik} - r_i\) for any potential match \((i, k) \in I \times K\).

Given a candidate set of equilibrium worker payoffs \(\{r_i\}\) combined with the i.i.d. Type 1 extreme value assumption for \(\epsilon_{ik}\) in equation (8), Decker et al. (2013) show that the probability that establishment \(m(k)\) wishes to fill position \(k\) by hiring (or continuing to employ) \(i\) is given by:

\[
P(i|k) = \frac{e^{\frac{\theta_g - r_i}{\sigma}}}{\sum_{i' \in I} e^{\frac{\theta_g' - r_{i'}}{\sigma}}} \tag{9}
\]

where I have suppressed the dependence of \(g\) and \(g'\) on \((i, k)\) and \((i', k)\), respectively. I can then use equation (9) to derive an expression for the probability that a randomly chosen position associated with destination type \(d\) wishes to hire a worker whose transition to the
position is classified into group \( g \):

\[
P(g|d) = \sum_{k \in d} P(g|d, k) P(k|d) = \frac{1}{|d|} \sum_{k \in d} P(g|k)
\]

\[
= \frac{1}{|d|} \sum_{k \in d} \sum_{i : g(i, k) = g} P(i|k)
\]

\[
= \frac{1}{|d|} \sum_{k \in d} \sum_{i : g(i, k) = g} \frac{e^{\frac{g_o - r_i}{\sigma}}}{\sum_{i' \in I} e^{\frac{g_o - r_{i'}}{\sigma}}}
\]

\[
= \frac{1}{|d|} \sum_{k \in d} \frac{\left( e^{\frac{g_o}{\sigma}} \right) \left( \sum_{i : g(i, k) = g} e^{-\frac{r_i}{\sigma}} \right)}{\sum_{i' \in I} e^{-\frac{r_{i'}}{\sigma}}},
\]

(10)

where \(|d|\) captures the number of positions \( k \) assigned to destination type \( d \). Next, I introduce two assumptions that allow one to express this conditional probability exclusively in terms of the group \( g \) and the types \( o \) and \( d \). First, Assumption 1 imposes that the mean exponentiated worker utility values \( e^{-\frac{r_i}{\sigma}} \) vary minimally across groups \( g \) featuring the same origin type \( o(g) \). Given the characteristics used to define \( o \) and \( g \) in the application below, this states that existing employees (potential stayers) and non-employees of each establishment have approximately the same mean value of \( r_i \) among workers whose initial jobs were in the same local area and pay category. In other words, the payoffs that workers in the same skill class require in equilibrium will not differ systematically across establishments within a small local area. This becomes a better approximation the more characteristics (such as occupation or education) are used to define an origin type \( o(i) \). To formalize this assumption, recall that the only characteristic \( z \) that distinguishes transition groups featuring the same combination of origin and destination types \((o,d)\) is an indicator for whether the worker \( i \) was already employed by \( k \) in the previous period, so that a given \((o,d)\) pair contains at most two groups, potential stayers and potential new hires. One can thus write:

Assumption 1: \[
\frac{1}{|g_k|} \sum_{i : g(i, k) = g} e^{-\frac{r_i}{\sigma}} \approx \frac{1}{|o|} \sum_{i : o(i) = o(g)} e^{-\frac{r_i}{\sigma}} = C_o(g) \forall (g, k) \tag{11}
\]

where \(|o|\) and \(|g_k|\) denote, respectively, the number of workers classified as origin type \( o \) and the number of workers whose transition would be classified as group \( g \) (either stayers or new hires among those in \( o \)) if hired by position \( k \) (a subset of the workers in \( o(g) \)). \( C_o \) denotes the mean value of \( e^{-\frac{r_i}{\sigma}} \) for a given origin group \( o \).

Second, Assumption 2 imposes that the share of potential stayers vs. potential new hires among workers from a given origin type \( o \) is common across establishments assigned to the same destination type \( d \). In the chosen context, this means that establishments in the same geographic area that are in the same industry supersector and same establishment size and establishment average pay categories have roughly the same number and skill composition
of employees. Let $S_{g|o,k}$ denote the share of workers of origin type $o$ who would be assigned to group $g$ if they filled position $k$ (i.e. potential stayers if $z(g) = 1$, movers if $z(g) = 0$), and let $\overline{S}_{g|o,d}$ be the mean of $S_{g|o,k}$ among all $k$ assigned to destination type $d$. Formally, I assume:

Assumption 2: $S_{g|o,k} \approx \overline{S}_{g|o,d} \forall k, \forall g$ \hfill (12)

Letting $f(o)$ denote the share of all workers assigned to origin type $o$, so that $|o| = f(o)I$, this implies:

$$|g_k| \equiv S_{g|o,k}f(o(g))I \approx \overline{S}_{g|o,d}f(o(g))I$$ \hfill (13)

These assumptions are necessary because the aggregate mean of $e^{-r_i}$, a non-linear function of a random variable, depends on its entire distribution. Essentially, the probability of filling a position with an existing employee depends on how many employees one already has, so that the group average depends on the establishment size distribution among firms who are at risk of creating a transition that could be classified into $g$. I am essentially hoping that Jensen’s inequality is close to equality ($f(E[X]) \approx E[f(X)]$) after conditioning on the characteristics that define the origin and destination types (most notably establishment size category).

Combined, these two assumptions imply that

$$\sum_{i:g(i,k)=g} e^{-r_i} \approx \overline{S}_{g|o,d}f(o(g))(I)C_o(g).$$ \hfill (14)

Applying these assumptions to the last expression in (10), I obtain:\footnote{Note that in contrast to Choo and Siow (2006), the probability that a worker of a given origin type $o$ is chosen by a position of destination type $d$ depends on share of workers of type $o$ in the population, $f(o)$. This difference stems from allowing an unobserved surplus component at the worker-position level. Menzel (2015) derives a similar formula in his nontransferable utility assignment model based on an unobserved surplus component at the agent pair level.}
\[
P(g|d) = \sum_{k \in d} \left( \frac{1}{|d|} \right) e^{\theta_g} \sum_{i: g(i,k)=g} e^{-r_i} \sum_{d' \in \mathcal{I}} e^{-r_{d'}}
\]
\[
= \sum_{k \in d} \left( \frac{1}{|d|} \right) e^{\theta_g} \sum_{g' \in (o,d)} \sum_{i: g(i,k)=g} e^{\theta_{g'} - r_{d'}}
\]
\[
= \sum_{k \in d} \left( \frac{1}{|d|} \right) e^{\theta_g} \sum_{g' \in (o,d)} S_{g|o,d} f(o)(I)C_o
\]
\[
= \frac{e^{\theta_g} S_{g|o,d} f(o)(I)C_o}{\sum_{g' \in (o,d)} e^{\theta_{g'}} \sum_{g' \in (o',d)} f(o')(I)C_{o'}}
\]

Let $\hat{\mu}$ denote an observed empirical matching. Since each observed transition can be assigned to a unique group $g$, one can easily aggregate the empirical matching into a group-level empirical distribution of transitions. Specifically, I let $\hat{P}_g$ denote the fraction of all observed transitions that are assigned to group $g$:
\[\hat{P}_g = \frac{1}{|d|} \sum_{i: 1(g(i,k) = g)} \]

Similarly, $\hat{f}(o)$ denotes the empirical fraction of all job transitions whose worker can be classified as type $o$:
\[\hat{f}(o) = \frac{1}{|d|} \sum_{i \in \mathcal{I}} 1(o(i) = o)\]

Finally, $\hat{h}(d)$ denotes the empirical fraction of all job transitions whose destination position can be classified as type $d$:
\[\hat{h}(d) = \frac{1}{|d|} \sum_{k \in \mathcal{K}} 1(d(k) = d)\]

Based on these definitions, one can estimate the (year-specific) conditional choice probability $P(g|d)$ by simply calculating the observed fraction of destination positions classified as type $d$ that were filled via transitions assigned to group $g$:
\[\hat{P}(g|d) = \frac{\hat{P}_g}{\hat{h}(d)}\]

Consequently, as the number of observed transitions gets large, each member of the set of empirical conditional choice probabilities $\{\hat{P}(g|d)\}$ should converge to the corresponding expression in (19). Note also that the average shares $\{\bar{S}_{g|o,d}\}$ can be assigned/estimated for each group $g$ using the average across all positions in group $g$ of the fraction of candidates for the position that are existing employees.

We are now ready to investigate the extent to which these assumptions, combined with the observed empirical choice probabilities $\{\hat{P}(g|d)\}$, can inform us about the mean match surplus values $\{\theta_g\}$. Consider the log odds between two conditional choice probabilities involving an (arbitrarily chosen) destination type $d_1$ and two (arbitrarily chosen) transition
Thus, we see that the adjusted log odds only identifies the relative mean (re-scaled) surplus values from transition groups $g_1$ and $g_2$, $\frac{\theta_{g_1} - \theta_{g_2}}{\sigma}$, in the case where both groups are associated with the same origin job type: $o(g_1) = o(g_2)$. Otherwise, the difference in surplus values is conflated with the log difference in mean exponentiated worker discounted utilities between the two origin job types $(\ln(C_{o(g_1)}) - \ln(C_{o(g_2)}))$.

However, now consider two additional transition groups $g_3$ and $g_4$ that are both associated with some destination type $d_2$ such that $o(g_3) = o(g_1)$ and $o(g_4) = o(g_2)$.\footnote{\textit{d}_2$ could be (but need not be) the same destination type as \textit{d}_1.} The four groups $g_1 - g_4$ can be chosen to represent the two pairs of transition groups that would be created by the two ways to match a given pair of destination positions to a given pair of workers. If one augments equation (21) by instead taking the log of the ratio of the (appropriately re-scaled) odds of $g_3$ and $g_4$ (conditional on $d_2$) and the (appropriately re-scaled) odds of $g_1$ and $g_2$ (conditional on $d_1$), one obtains:

\[
\ln\left(\frac{\hat{P}_{g_3|d_2}/(\overline{s}_{g_3|o(g_3),d_2},f(o(g_3)))}{\hat{P}_{g_4|d_2}/(\overline{s}_{g_4|o(g_4),d_2},f(o(g_4)))}\right) / \ln\left(\frac{\hat{P}_{g_1|d_1}/(\overline{s}_{g_1|o(g_1),d_1},f(o(g_1)))}{\hat{P}_{g_2|d_1}/(\overline{s}_{g_2|o(g_2),d_1},f(o(g_2)))}\right)
\]

\[= \left(\frac{\theta_{g_3} - \theta_{g_1}}{\sigma}\right) - \left(\theta_{g_1} - \theta_{g_2}\right) - \left(\frac{\theta_{g_4} - \theta_{g_1}}{\sigma}\right) + \left(\ln(C_{o(g_3)}) - \ln(C_{o(g_4)})\right) - \left(\frac{\theta_{g_1} - \theta_{g_2}}{\sigma}\right) - \left(\ln(C_{o(g_1)}) - \ln(C_{o(g_2)})\right)
\]

\[= \left(\frac{\theta_{g_3} - \theta_{g_1}}{\sigma}\right) - \left(\theta_{g_1} - \theta_{g_2}\right)
\]

Thus, we see that the appropriate log odds ratio can identify the expected gain in mean
scaled surplus values from a swap among partners from any two end-of-year job matches. Note that these difference-in-differences do not preserve information about the baseline welfare of either worker types or position types: the mean discounted value of each worker type and each destination position type gets eliminated during the differencing/conditioning, respectively.

However, the set of difference-in-differences $\Theta^D \equiv \{ (\theta_g - \theta_{g'} - (\theta_{g''} - \theta_{g'''}) \sigma : o(g) = o(g''), o(g') = o(g'''), d(g) = d(g'), d(g'') = d(g''') \}$ preserves the crucial information about the relative efficiency of different matchings that exists in the observed transition group frequencies.

Specifically, in the next subsection I show that identification of the set of surplus difference-in-differences is sufficient to generate the unique assignment in counterfactual simulations that forecast the aggregate distribution of transition types $P(g)$ for any arbitrary change in either the marginal distribution of worker origin match types $f(o)$ or the marginal distribution of destination position types $h(d)$ (or both). Furthermore, if more than one observed matching is available, then $\sigma$ can potentially be (roughly) estimated as well, allowing for a proper welfare analysis that calculates the approximate mean utility and profit gain for each worker origin type and position destination type, respectively, from any such shifts in labor supply or demand.

3.2 Counterfactual Simulations

This subsection demonstrates how to predict the way in which a set of workers (initially matched with a set of positions) would be reallocated to a new set of positions, given a particular job matching technology (i.e. collection of production functions, utility functions, and search and recruiting cost functions). In the empirical work, such counterfactuals will involve altering the distribution of destination positions by introducing labor demand shocks of various forms.

One can characterize the set of workers to be reallocated using their distribution across origin match types, $f^{CF}(o)$. The “CF” superscript indicates that this could potentially be a counterfactual distribution (e.g. capturing a proposed influx of refugees). Similarly, the set of counterfactual positions to be filled can be represented by its type distribution $h^{CF}(d)$, and the prevailing technology can be denoted $\{\theta^{CF}_g\}$. The values $f^{CF}(o)$, $h^{CF}(d)$, and $\{\theta^{CF}_g\}$ are all treated as inputs that are either observed or constructed by the researcher/policymaker. The goal is to use these inputs to predict the equilibrium distribution of transitions across transition groups, $\{g = 1, \ldots, G\}$, as captured by $P^{CF}(g)$.

Consider the case of a manufacturing plant considering relocation. The immediate change in the location of a set of manufacturing and management positions that would occur is known by a local development board, and the existing group mean surpluses $\{\theta^{CF}_g\}$ have been estimated; the board wishes to predict the extent to which the plant relocation
will decrease the probability of nonemployment and more generally increase the job-related utility among existing workers/job seekers in the local area versus workers arriving from neighboring or distant locales (and perhaps the profits of other local firms relative to more distant firms).

I assume that the counterfactual assignment also satisfies Assumptions 1 and 2 above, which implies that

$$\sum_{i: g(i,k) = g} e^{-\epsilon_{ik}} \approx \mathcal{P}^{CF}(y_o(g),d(g)) f^{CF}(o(g)) I \mathcal{C}_o^{CF}$$

for sets of group- and type-specific constants \( \{S_{g|o,d}\} \) and \( \{C_o^{CF}\} \). I also assume that the average share of workers of each origin type who are existing employees of a random position in each destination type, \( \{S_{g|o,d}\} \), is known, and treat it as an input. In particular, when \( f^{CF}(o) = f^{y'}(o) \) and \( h^{CF}(d) = h^{y'}(d) \) for some observed year \( y' \), then the appropriate existing employee fractions can be obtained via \( S_{g|o,d} = S_{y'|o,d} \forall g \), which is observed. Maintaining the assumed Type 1 distribution for the idiosyncratic component of the match surplus \( \epsilon_{ik} \), the counterfactual conditional choice probability \( \mathcal{P}^{CF}(g|d) \) can be expressed as:

$$\mathcal{P}^{CF}(g|d) = \frac{e^{\theta_{g CF}} \mathcal{S}_{g|o(d)}^{CF} f^{CF}(o(g)) \mathcal{C}_o^{CF}}{\sum_{o' \in \mathcal{O}} \sum_{g' \in (o, d)} e^{\theta_{g CF}} \mathcal{S}_{g'|o(g')}^{CF} f^{CF}(o') \mathcal{C}_{o'}^{CF}}$$  \hspace{1cm} (23)

The origin type-specific mean worker discounted utility values \( \{C_1^{CF} \ldots C_O^{CF}\} \) are equilibrium objects that will be affected by the counterfactual changes in technology incorporated into \( \{\theta_g^{CF}\} \) and the counterfactual changes in the composition of both supply and demand unknown incorporated into \( f^{CF}(o) \) and \( h^{CF}(d) \), and are thus unknown. As a result, each counterfactual conditional choice probability cannot be immediately constructed, and thus must be treated as a function of the vector of type-specific mean worker exponentiated discounted utility values.

Galichon and Salanié (2015) and Decker et al. (2013) each show that the probability distribution over transition groups \( \mathcal{P}^{CF}(g) \) that satisfies the aggregate analogues to the stability and feasibility conditions is unique. Galichon and Salanié (2015) show that this stable matching and the corresponding mean payoffs of types on each side of the market can be computed by an interative projection fitting procedure.

However, the number of unmatched partners on both sides of the market are assumed to be observed when proving identification and in constructing counterfactual assignments in each of Choo and Siow (2006), Decker et al. (2013), and Galichon and Salanié (2015). While I can (roughly) observe the number of unemployed workers of each type, the LEHD database described below contains no information about unfilled vacancies. Observing only the subset of positions \( \tilde{K} \in \mathcal{K} \) that match with workers does not pose major problems for identification of the relative joint surplus parameters, because each subset of assignments within a stable matching must also be stable. Thus, stability among the matrix of observed year-to-year transitions between dominant jobs based on \( \mathcal{I} \) and \( \tilde{K} \) is a necessary condition for stability of
the full market transition matrix defined by $I$ and $K$. Therefore, the relationships between the transition surpluses $\{\pi_{ijk}\}$ that I recover would not be reversed if data were augmented with additional transitions and unmatched agents.

In principal, though, unfilled positions may put upward pressure on wages that affect the division of surplus between workers and positions, even if they do not affect the final assignment of workers to positions. Furthermore, unfilled positions might potentially become filled in the wake of the potential labor supply and demand shocks that I simulate, so that the incidence measured in the counterfactual simulations described below may be slightly distorted. For example, a tornado that eliminates a number of local jobs may depress local wage levels far enough for previously unfilled positions to become the most desirable options for some local workers.

To formally rule out such scenarios, I assume that no position captured by $h^{CF}(d)$ will choose to remain vacant in any the counterfactual equilibria I seek, and that any vacancies that are not reflected in $h^{CF}(d)$ will remain vacant in such equilibria and furthermore will never the second-best option for any worker who takes a job in the destination period in the counterfactual simulations. This assumption implies that the unfilled positions do not affect either the allocation of workers to positions nor the division of surplus among them. But it also shuts down extensive margin adjustments of the number of positions demanded by establishments in response to changes in wages (unless they are built in to $h^{CF}(d)$ ex ante), so that establishments can only adjust the composition of workers they choose to fill a fixed set of positions. Violations of this perfect inelasticity assumption will likely lead me to overstate the magnitude of welfare gains from positive shocks and losses from negative shocks, since some establishments might reduce the number of positions they seek to fill when increased local labor demand increases the equilibrium payoffs positions must offer to workers, and some previously establishments might create new positions when decreased competition lowers the required worker payoffs. If there are sizeable adjustment costs to changing the number of positions at establishments, and the changes in the cost of an efficiency unit of labor from the relatively small and geographically targeted shocks I consider are sufficiently small, such violations need not produce appreciable bias in incidence forecasts.

While strong, the perfect inelasticity assumption dramatically simplifies the otherwise demanding computation of counterfactual equilibria. With an unknown number of unmatched partners on each side, Galichon and Salanié (2015) show that one must solve a set of $O + D$ non-linear equations that combine the feasibility and stability conditions for the

---

19 Instead, positions who eventually settle for other workers represent the best outside option for each worker, in the following sense: for each worker, there exists a position that ends up being filled by another worker that could increase its payoff by hiring the chosen worker at a wage yielding a higher utility to the chosen worker than their utility at any wage any unfilled position would prefer to remaining unfilled. (Of course, the filled position cannot increase its payoff by outbidding the position that actually hires the chosen worker, since the observed matching is presumed to be stable).
mean equilibrium payoffs of all worker types \( \{C^C_0 \} \) and firm types (denoted \( \{C^F_d \} \)). By contrast, when unfilled vacancies are ruled out, the number of positions of each destination type that are “supplied” is known, and must in equilibrium be equal to the number of such positions demanded by workers. This implies a set of \( D \) type-level market clearing conditions that can be used to solve for \( \{C^F_d \} \). \(^{20}\) Equivalently, if an extra, dummy “position” type is added with mass equal to the share of workers who will be left unmatched (knowable in advance given assumed knowledge of the number of filled positions), then the augmented “demand” (including “demand” from nonemployment) for each origin type \( o \) will equal supply for each type \( o \) (given by \( f^C_F(o) \)), so that one can clear the market on the worker side.\(^{21}\)

Since 1) only relative payoffs required by alternative origin types matter in determining the equilibrium assignment (so that one can normalize \( C^C_1 = 0 \)), and 2) the origin-type distribution \( f^C_F(\ast) \) must sum to one, one obtains the following \( O - 1 \) market clearing conditions:

\[
\sum_{d \in D} h^C_F(d) \left( \sum_{g:o(g)=2} P^C_F(g|d, C^{CF}) \right) = f^C_F(2) \\
\vdots \\
\sum_{d \in D} h^C_F(d) \left( \sum_{g:o(g)=0} P^C_F(g|d, C^{CF}) \right) = f^C_F(O)
\]

(24)

These market clearing conditions represent a system of \( O - 1 \) equations with \( O - 1 \) unknowns, \( C^{CF} \equiv \{C^C_2, \ldots, C^C_O \} \). Given a solution to this system, one can then construct the counterfactual transition probability for any transition group via \( P^C_F(g) = \sum_d h^C_F(d) P^C_F(g|d, C^{CF}) \). Note that the inability to observe unfilled vacancies prevents the determination of the changes in absolute levels of worker utility following labor supply and demand shocks.

Since any solution to this system also satisfies the stability and feasibility conditions, it must be the unique aggregate counterfactual equilibrium assignment. This line of reasoning reveals that one can construct a unique counterfactual transition between labor market

\(^{20}\) Koopmans and Beckmann (1957) point out that when unmatched agents only exist on one side of the market, the dual problem payoffs need only be recovered on one side of the market in order construct the stable assignment.

\(^{21}\) These dummy nonemployment positions represent a computational mechanism for appropriately incorporating the mean surpluses workers obtain from nonemployment, \( \{\pi^0_o\} \). I formally prove this result as Proposition A1 in Appendix A1. The intuition behind the proof is that the unique assignment could have been computed by aggregating the supply-side stability conditions \( \mu_k = 1 \text{ iff } k \in \arg \max_{k \in \tilde{K} \cup \{0\}} \pi_{ik} - q_k \) instead. This can be motivated via an assignment mechanism where workers bid for positions instead of the reverse. But under the assumption that all positions will be filled, “demand” for positions of each destination type \( d \) will exactly equal supply. I then show that the group-level assignment that satisfies the resulting system of excess-demand equations (and is thus the unique group-level assignment consistent with an individual-level stable matching) will also satisfy the corresponding system of equations for the worker-side of the market, provided that the appropriate “demand” from dummy nonemployment positions is added.
allocations given any marginal distributions of origin and destination job matches \((f^{CF}(\ast)\) and \(h^{CF}(\ast)\)) and any vector of mean group surplus values \(\{\theta^{CF}_g\}\) (and any vector of existing employee shares \(\{S^{CF}_{g|o,d}\}\)).

Because this approach only requires solving \(\min\{O, D\}\) non-linear equations, it provides considerable computational savings when the number of types is much larger on one side of the market than the other. Below I will present results that average over 500 counterfactual allocations, each featuring around 1,000 origin types and around 10,000 destination types, that would be prohibitive to compute if unmatched agents on both sides were permitted.

### 3.3 Interpreting the Counterfactual Simulations

In the counterfactual labor demand shock simulations below, I will generally use data from the 2010-2011 set of job transitions (including stayers) to construct the simulation inputs, so that \(\Theta^{CF} = \Theta^{2010}\), \(f^{CF}(\ast) = f^{2010}(\ast)\), and \(h^{CF}(\ast)\) will equal \(h^{CF}(\ast)\) plus a shock consisting of additional positions added to or subtracted from a particular destination type \(d\) (or relocated from one type \(d'\) to another type \(d\)). I wish to interpret the difference between the resulting counterfactual reallocation and the reallocation that actually occurred between 2010 and 2011 as the one-year impact that such a stimulus, disaster, or plant relocation would have caused in that economy. However, a few additional assumptions and clarifications are necessary to justify and elaborate on this desired interpretation.

First, constructing the system of market-clearing conditions (24) requires providing the full set of 2010-2011 group-level joint surpluses \(\Theta^{2010} \equiv \{\theta^{2010}_g\}_{g \in G}\), despite the fact that the identification argument in section 3 suggests that only the set of surplus difference-in-differences \(\Theta^{D-in-D,2010}\) is identified. In Appendix A2, I prove the following proposition:

**Proposition 1:**

Define the set \(\Theta^{D-in-D} \equiv \{(\theta_g-\theta_{g'})-(\theta_{g''}-\theta_{g'''}): o(g) = o(g''), o(g') = o(g'''), d(g) = d(g''), d(g') = d(g''')\} \cap \{g \in G\}\). Given knowledge of \(\Theta^{D-in-D}\), a set \(\tilde{\Theta} = \{\tilde{\theta}_g \forall g \in G\}\) can be constructed such that the unique group level assignment \(P^{CF}(g)\) that satisfies the system of excess demand equations (24) using \(\theta^{CF}_g = \tilde{\theta}_g \forall g \in G\) and arbitrary marginal PMFs for origin and destination types \(f^{CF}(\ast)\) and \(g^{CF}(\ast)\) will also satisfy the corresponding system of excess demand equations using \(\theta^{CF}_g = \theta_g \forall g \in G\) and the same arbitrary PMFs \(f^{CF}(\ast)\) and \(g^{CF}(\ast)\). Furthermore, denote by \(\tilde{C}^{CF} \equiv \{\tilde{C}_1^{CF}, \ldots, \tilde{C}_O^{CF}\}\) the market-clearing utility values that clear the market using \(\theta^{CF}_g = \tilde{\theta}_g\), and denote by \(C^{CF} \equiv \{C_1^{CF}, \ldots, C_O^{CF}\}\) the market-clearing utility values that clear the market using \(\theta^{CF}_g = \theta_g\). Then \(\tilde{C}^{CF}\) will satisfy \(\tilde{C}_o^{CF} = C_o^{CF} e^\frac{-\Delta_o}{\sigma} \forall o \in O\) for some set of origin type-specific constants \(\{\Delta_o\}\) that is invariant to the choice of \(f^{CF}(\ast)\) and \(g^{CF}(\ast)\).
Essentially, the proposition states that the identified set of surplus difference-in-differences $\Theta^{D-in-D}$ contains sufficient information to generate the unique counterfactual group-level assignment $P^{CF}(g)$ that would be consistent with the corresponding true set of surpluses $\Theta$. Furthermore, the vector of utility premia $\tilde{C}^{CF}$ that clears the market using the artificial surpluses $\tilde{\Theta}$ generated from the surplus difference-in-differences will always differ from the “true” premia $C^{CF}$ that clear the counterfactual market under $\Theta$ by the same $\sigma$-type-specific constants regardless of the compositions of supply $f^{CF}(o)$ and $h^{CF}(d)$ used to define the counterfactual.

While absolute levels of counterfactual mean utility by origin type are never uniquely determined (even when $\Theta$ is fully known), the existence of the “bias” terms $\{\Delta_o\}$ in Proposition 1 indicates that the relative levels of utility among origin types in counterfactual allocations (including the true “counterfactual” that was observed) are not identified. One cannot infer the level of relative utility among workers at different skill levels who start the time period in different locations. This inability to determine the existing division of surplus among workers and firms for any origin and destination type combination, which does not appear in Choo and Siow (2006) or Galichon and Salanié (2015), stems from the assumptions necessary to accommodate the lack of data on unfilled vacancies.

However, because the “bias” terms $\{\Delta_o\}$ are constant across counterfactuals featuring different supply and demand compositions $f^{CF}(o)$ and $h^{CF}(d)$, the relative differences in origin-type mean utilities, $[[\ln(C^{CF}_o^1) - \ln(C^{CF}_o^2)) - (\ln(C^{CF}_o^1) - \ln(C^{CF}_o^2))] \approx (\frac{\sigma}{\sigma})$, among two counterfactuals across origin types can be identified.\(^{22}\) Note that such a pair of counterfactuals might include one that features a stimulus package or natural disaster versus an otherwise identical counterfactual that does not. For certain counterfactual scenarios, one might plausible restrict utility gains for a particular origin type reference group to be known, thereby allowing (scaled) utility gains or losses $\frac{\sigma}{\sigma}$ for other origin-type groups to be identified. Below, I assume that a plant relocation generates zero aggregate utility gains for workers (since aggregate labor demand did not change), creating a natural reference group (the population-weighted mean across all origin types) for whom the change is assumed to be zero. Alternatively, I assume that the small, very local stimuli and natural disasters I consider generate zero utility gain or loss for workers on the opposite side of the country. Such restrictions allow the share of total welfare gains (or losses) from a labor demand (and/or supply) shock experienced that redound to each origin worker type to be determined. Furthermore, because the model is symmetric with respect to workers and positions, this also implies that the mean changes in destination position-type discounted profits $\frac{\sigma}{\sigma}$ can also be identified, so that the share of profit gains (or losses) can also be computed. Thus, when combined with the available data, the model permits a reasonably complete analysis of job-related welfare incidence from labor

\(^{22}\)This approximation requires that the variation in utility values among the workers assigned to the same origin type is limited, so that $\ln(C_o) \equiv \ln(\frac{1}{|o|} \sum_{i:o(i)=o} e^{\frac{r_o}{\sigma}}) \approx \ln(\frac{1}{|o|} \sum_{i:o(i)=o} e^{\frac{r_o}{\sigma}})^{\frac{\sigma}{\sigma}}$. 26
supply and demand shocks.

Second, in addition to these necessary normalizations, in order for the counterfactual allocation and predicted welfare gains to accurately represent what would have happened had the simulated shocks occurred, one must also assume that the set of joint surpluses $\Theta^{2010}$ and the destination type marginal distribution $h^{2010}(\ast)$ that serve as inputs to the simulation are exogenous to (i.e. unaffected by) the shock itself. Any reallocation and welfare changes are assumed to be driven exclusively by the change in market clearing transfers across origin types required to eliminate the imbalances between supply and demand that the shock generates.

This exogeneity assumption would almost certainly not hold exactly if such shocks had actually occurred. Specifically, it requires that the labor demand shock does not induce further changes in either firms’ location decisions or the number of positions they wish to fill. To the extent that establishment relocations or startups cause other firms to form, relocate, or expand, these additional compositional changes would need to be anticipated by the researcher and incorporated into the destination type distribution $h^{CF}(\ast)$ provided to the model to ensure that the true net change in destination type composition is captured, perhaps exploiting estimates from other spatial models that focus on product markets. One hopes that the shocks considered are sufficiently modest in scale that incumbent firms’ location and expansion/contraction decisions would not have been substantial enough to qualitatively change the estimated impact.

Along the same lines, there are plausible mechanisms by which the joint surplus parameters $\Theta^{2010}$ might respond to the shock. First, to the extent that production agglomeration economies exist, the existence of a new establishment nearby might increase the demand for intermediate products produced by other local firms, thereby raising the productivity of workers for such firms. Second, if the new jobs are thought to be persistent and search/recruiting/moving costs increase with distance, then obtaining a job at a different establishment in the same local area as a newly relocated establishment might now have greater continuation value because future job searches will begin in a local area featuring a higher level of labor demand. Again, one could argue that such surplus changes might be minuscule relative to the size of other surplus determinants that are plausibly invariant to the shock, such as how productive each worker type is at each firm type, how much different worker types value firm and location amenities, and moving costs from alternative transitions, so that such exogeneity violations would generate minimal bias.

Furthermore, while both of these scenarios change the joint surpluses $\theta_g$ of transition groups involving nearby firms, if the increased productivity of workers or the increased continuation value for workers at such firms is common to all potential origin types $o$, then it will not affect the surplus difference-in-differences $\Theta^{D-in-D}$ that generate the counterfactual assignment. Furthermore, the proof of Proposition 1 in Appendix A2 shows that such composition-induced surplus changes also will not affect the equilibrium origin type payoffs
that capture the shock’s incidence among groups of workers.\textsuperscript{23} Instead, any increase in joint surplus by a destination position $k$ that is common to all workers will be fully reflected in $k$’s profit payoff, either through higher revenue for the same costs (agglomeration case) or through lower salaries that offset the change in worker continuation value.\textsuperscript{24} Thus, bias in forecasted worker incidence in counterfactuals from shock-induced changes in joint surpluses only stems from differential changes in joint surplus among origin worker types for a given destination type.

Another caveat relates to the permanence of the shock. I restrict attention to forecasting reallocations and welfare changes that occur within one year of the chosen shocks, and for the stimulus and relocation packages I generally assume that the new positions generate the same surplus values $\theta_g$ as existing positions of the chosen destination type. Implicitly, this requires that the new positions have the same expected duration over time as any other position of their type. If one wished to simulate a particularly temporary construction stimulus, one would need to estimate a separate set of surplus parameters for temporary versus permanent construction jobs.\textsuperscript{25} The model nonetheless highlights the degree to which the incidence of even very local labor demand shocks is likely to be spread quite widely across space (and across skill levels) even over a short time horizon, despite the fact that short-run mobility frictions may be quite large (a large share of job transitions occur among very nearby locations, as documented below).

A final, important caveat relates to the absence of a housing market in the model (and the corresponding absence of residential choices in the data). Standard models of spatial equilibrium in urban economics (e.g. Roback (1982) or Kline and Moretti (2013)) emphasize the critical role of the housing market in determining incidence from place-based policies. In particular, if housing supply is quite inelastic and workers are very mobile, a large share of the incidence of a positive place-based shock is enjoyed by landholders in the form of higher rents (which offset the utility gains to workers from any wage increases). Thus, in principle failing to model the housing market could result in meaningfully biased estimates of overall shock incidence. However, this same argument also suggests that for local workers who are also nearby renters, the simulated employment-related welfare gains are likely to place

\textsuperscript{23}Using the proof’s notation, such surplus changes will only change $\Delta_2$, which does not enter into equilibrium mean payoffs for origin types $\{C_{CF}^o\}$.

\textsuperscript{24}Note, though, that in these scenarios the profit gains among nearby position types $d$ will be understated. The possibility of differential agglomeration effects for nearby firms across different shock compositions (emphasized by Glaeser et al. (2008)) is one reason that I focus primarily on incidence among workers, for whom differential agglomeration effects are likely to be of second order importance.

\textsuperscript{25}More generally, a more precise distinction of differences in welfare effects between shocks of different expected durations requires a fully dynamic assignment model along the lines of Choo (2015) in which worker expectations and the serial correlation in (now time dependent) idiosyncratic surplus components $\epsilon_ikt$ would need to be specified. Similarly, the Markov-style model used here conditions only on a single previous period’s location, and thus cannot accommodate optimal sequences of location decisions and return migration choices highlighted by Kennan and Walker (2011) that would enrich a dynamic analysis. See Weinstein (2018) for an example of an attempt to evaluate the dynamic consequences of a particular local shock.
an upper bound on the full welfare gains these workers would experience. If certain types of local policies focused on bringing “good” jobs to town generate an employment-related incidence that is either concentrated among higher skilled workers or is quite geographically dispersed among workers in many distant locations, then local low-skill workers would be well justified in resisting such local development initiatives.\textsuperscript{26}

Furthermore, in contexts where housing supply is likely to be relatively elastic (such as rural areas or areas with weak zoning laws) or where there exists excess housing supply due to a declining population, housing prices may move little, and abstracting attention from the housing market may produce little bias in incidence forecasts. Indeed, Gregory (2013) finds that neighborhoods receiving empowerment zone status, a local labor demand shock similar to those I estimate, experienced negligible changes in rent but substantial wage gains among residents, suggesting that omitting a housing market response might generate minimal bias.\textsuperscript{27}

Along the same lines, in contexts where tastes for particular neighborhoods is strong but commuting between neighborhoods is fairly low cost, much of the adjustment to shocks may take the form of changing commuting patterns, with very little change in demand for housing across locations. Indeed, since commuting costs from job transitions that involve locational changes already implicitly constitute a component of the joint surplus $\theta_g$, they are appropriately captured by the model; thus a scenario in which the true response to a labor demand shock involves lots of small commuting adjustments that ripple out from the focal point of the shock is likely to be closely matched by the model-based simulations, with potentially accurate incidence estimates.\textsuperscript{28} While a full welfare analysis certainly requires incorporating housing and product markets, the goal of the approach taken here is to highlight the heretofore underappreciated role of differential geographic scopes of local labor markets for different types of workers and firms in determining the incidence of alternative local labor demand interventions.

### 3.4 Identifying $\sigma$

While the share of welfare gains or losses for workers (or firms) can be identified without knowledge of $\sigma$, $\sigma$ is nonetheless a parameter of consider interest. Under the assumptions that the worker and position payoffs are additive in worker earnings, knowledge of $\sigma$ would

\textsuperscript{26}Exceptions to this claim might occur, for example, if house price increases yielded property tax revenue that was disproportionately spent on services these workers/residents valued.

\textsuperscript{27}The authors note that the neighborhoods receiving these shocks were often in locations experiencing recent decline that were unattractive residential options for many, so that the lack of impact on local rent rates may not generalize to shocks to healthier locations.

\textsuperscript{28}Note also that mobility frictions induced by housing markets are likely to be partly reflected in the log odds ratios capturing the relative propensities with which different origin worker types make certain types of job transitions that are used to identify the set $\Theta^{O-m-D}$. So differential willingness to pay high prices for locational amenities will be captured by heterogeneity in $\{\theta_g\}$ across origin worker types for groups involving positions in the same location, and thus will be reflected in the counterfactual simulations.
allow the estimated scaled utility gains $\frac{\pi_{CF_1} - \pi_{CF_2}}{\sigma}$ for origin worker types and scaled profit gains $\frac{\pi_{CF_1} - \pi_{CF_2}}{\sigma}$ to be re-scaled in dollar terms, making it easy to gauge whether the utility gains or losses from a given labor market shock are economically meaningful.

Recall that $\sigma$ captures the relative contribution of the idiosyncratic worker-position components $\epsilon_{ij(k)}$ versus systematic group-level components $\theta_g$ to the overall variance in joint surpluses across all worker-position pairs. Intuitively, when one observes a given destination position type $C$ choosing an origin worker type $A$ much more often than origin worker type $B$, it could be because $\theta_{AC} \gg \theta_{BC}$ even though $\sigma$ is moderately large, or because $\theta_{AC}$ is marginally larger than $\theta_{BC}$ but $\sigma$ is tiny. If the former is true, clearing the market after a shift in destination position composition could require large changes in the utility values that must be promised to workers to engender sufficient substitution across worker types to overcome strong comparative advantages from matching certain types of workers and positions together. If the latter is true, very small utility changes would suffice. Thus, $\sigma$ determines the elasticity of worker choices with respect to relative wages, and is rather important in determining the degree to which changes in labor demand composition cause substantial reallocation of utility across skill types and geographic areas.\(^{29}\)

As Galichon et al. (2017) have noted, $\sigma$ is not identified from a single observed matching. However, combining information from multiple matchings can potentially identify $\sigma$. Since I observe national market-level matchings (transitions from a set of origin job matches to a set of destination job matches) for each pair of years 2003-2004 to 2010-2011, I attempt to calibrate a value of $\sigma$.

I exploit the fact that the composition of U.S. origin and destination job matches $f^y(o)$ and $h^y(d)$ evolved across years $y$. Specifically, I estimate the set of group-level surpluses $\{\theta^2002_g\}$ from the observed 2002-2003 matching (the first year where all 19 states in the sample provide data). Then, holding these surplus values fixed, I combine $\{\theta^2002_g\}$ with $f^y(o)$ and $h^y(d)$ from each other year $y \in [2003, 2010]$ to generate counterfactual assignments and changes in mean (exponentiated) scaled utility values $\{C_{CF,y}^o\}$ for each origin type. These counterfactuals predict how mean worker utilities by skill/location combination could have been expected to evolve over the relevant period given the observed compositional changes

---

\(^{29}\)Note that the existence of a single $\sigma$ parameter governing relative wage elasticities is driven by assuming additive separability of the mean choice-specific values and an i.i.d type 1 extreme value distribution for the vector of idiosyncratic surpluses for particular worker/destination position combinations. As shown by Menzel (2015), the identification argument and the counterfactual simulations are not sensitive to the assumption of a type 1 extreme value distribution per se. Instead, the key assumption is that the idiosyncratic worker/position surpluses are independently and identically distributed across all alternative worker/position matches, which leads choices by one side of the market to satisfy the independence of irrelevant alternatives (IIA) property conditional on required values to be offered the other side. As the relevant matching market gets large (as the U.S. national labor market most certainly is), the counterfactual allocation associated with any joint distribution of idiosyncratic surpluses satisfying this property will converge to a common, unique allocation. Galichon and Salanié (2015) and Chiappori et al. (2009) discuss the possibility of relaxing the i.i.d assumption to allow certain forms of correlation in the idiosyncratic component across matches featuring shared characteristics. However, I maintain the standard i.i.d. assumption in this paper in order to minimize an already substantial computational burden.
in labor supply and demand had the underlying surplus values $\Theta$ been constant and equal to $\Theta_{2002}$ throughout the period.

I then regress differences in the actual mean annual earnings changes experienced among different origin types $o$ from origin periods $y-1$ to destination periods $y$ on the corresponding log differences in predicted changes in mean scaled utility values $\ln(C_{o}^{CF,y}) - \ln(C_{o}^{CF,y'})$, which approximately equal $\frac{r_{o}^{CF,y}-r_{o}^{CF,y'}}{\sigma_{y}}$ if individual-level utility changes $r_{i}^{CF}$ are roughly similar within each origin type $o$. To the extent that a) most of evolution in the utility premia enjoyed by workers in particular locations and skill categories was due primarily to changes in supply and demand composition rather than changes in the moving costs, recruiting costs, tastes, and relative productivities that compose the joint surplus values $\Theta$, and b) mean utility gains for each origin type generally came from increases in mean annual earnings in the chosen year rather than increases in non-wage amenities or continuation values, the coefficient on $\frac{r_{o}^{CF,y}}{\sigma_{y}}$ will approximately equal $\sigma_{y}$. Appendix A3 provides further detail on this procedure.

Clearly, given the additional strong assumptions required, this approach represents a relatively crude attempt to calibrate $\sigma$. In practice, the estimates I obtain for $\sigma$ are surprisingly consistent across years. I used the mean estimate of $\sigma_{y}$ across all sample years, $\bar{\sigma} = 19,600$, to produce dollar values for all the results relating to utility gains presented below.

As noted by Galichon and Salanié (2015), when match-level unobserved heterogeneity is assumed away as in Choo and Siow (2006), the observed earnings of the matches between origin workers and destination positions can be further used to attempt to decompose each group-level joint surplus value $\theta_{g}$ into its worker and position subcomponents (denoted $\theta_{l}^{g}$ and $\theta_{f}^{g}$ respectively). In other words, one can determine the relative contributions of amenities/future earnings opportunities versus current and future revenue contributions to $\theta_{g}$. However, in Appendix A4 I show that clean identification of $\theta_{l}^{g}$ and $\theta_{f}^{g}$ breaks down without the particular structure Choo and Siow (2006) place on the unobserved match component $\epsilon_{ik}$ unless further strong assumptions are imposed. I chose not to pursue this path further in this paper, primarily because I have shown that this decomposition is unnecessary to determine the incidence across worker and position types of alternative local labor demand shocks, which is the primary goal of the paper.

4 Data

4.1 Overview

I construct a dataset of year-to-year worker job transitions (pairs of primary jobs in consecutive years) using the Longitudinal Employer-Household Dynamics (LEHD) database. The core of the LEHD consists of state-level wage records collected for unemployment insurance
purposes that contain quarterly worker earnings and unique worker and establishment IDs for a near universe of jobs in the state. The worker and establishment IDs are then linked across states, and the data are augmented with information on firm- and establishment-level characteristics (notably establishment locations and detailed industry codes) from a state-supplied extract of the ES-202/QCEW report and individual-level data from the Social Security Administration (including age, race and sex but not including occupation nor education for most of the sample).

4.2 Sample Selection

My sample consists of all LEHD records from the 19 U.S. states that agreed to provide data to my FSRDC project. By agreement with the Census Disclosure Avoidance Review, the identities of the states cannot be revealed, but they include large, medium, and small states, and are spread throughout the U.S., albeit unevenly. Different states begin reporting data in different years, with some states reporting as early as 1990 and others entering as late as 2003. For the model validation exercise described in Section 6.3 below, I consider any actual shock taking place in one of the target states after 1997. All of the counterfactual shocks, however, are based on model parameters estimated from 2010-2011 data. Preliminary work suggested that the forecasts of shock incidence were quite insensitive to the pair of years used to identify the surplus parameters.

I also restrict the sample to person-years featuring individuals with ages between 20 and 70. This restriction limits the influence of “nonemployment” spells consisting of full-time education or retirement followed by part-time work, so that parameters governing nonemployed workers would be identified primarily from prime-aged workers who were unemployed or temporarily out of the labor force.

The initial dataset is converted from a job-quarter-year-level dataset to one whose observation level is the combination of a person and a pair of primary jobs in consecutive years (i.e. person-level job transition or retention). This is done by first identifying each individual’s primary job in each year, then aggregating earnings from the primary job across all quarters within the year, and then appending the primary job from the following year to the current observation to create a transition/retention observation. The primary job for a worker is defined as the job with the highest earnings that exists for at least one full quarter. A worker who does not report earnings above $2,000 at any job in any full-quarter in

---

30 The database does not include farm jobs or self-employed workers. I also exclude federal employees, who must be merged in via a separate OMB database. This has little consequence given that the current sample does not contain Virginia, Maryland, and the District of Columbia.

31 For further details about the contents and construction of the LEHD, see Abowd et al. (2009).

32 This was true despite the trend toward decreasing worker job-to-job mobility over this time period documented by Hyatt et al. (2016).

33 For computational reasons, the results in this draft are based on a 50% random sample of all transition-level observations from the sample just defined.

34 A job is observed in a full quarter if the worker-establishment pair reports positive earnings in the
a given year in any observed state (including observed states outside of the chosen sample of states for the more recent years) is designated nonemployed. An individual is included in the sample if he/she is ever observed as employed in one of the sample states in any of the years for which data was provided.

I limit each sample member’s years in the sample to consist of years between his/her first and last years of observed employment. Thus, the sample only includes spells of nonemployment that are bookended by spells of observed employment. I exclude nonemployment spells before the first year and after the last year of observed employment in an attempt to remove workers who are out of the labor force.

Observing only employment among the limited sample of states providing data creates two problems that must be addressed. First, since workers who are not observed working in the sample states are initially classified as nonemployed, many spurious employment-to-unemployment (denoted E-to-U) transitions, unemployment-to-unemployment (U-to-U), and unemployment-to-employment (U-to-E) transitions will be generated in which the “unemployed” worker was actually employed in an out-of-sample state. I address this by using data from the American Community Survey to construct estimated counts of true U-to-E, E-to-U, and U-to-U transitions for each combination of origin U.S. state, destination U.S. state, age category (10 brackets), destination industry supersector (for U-to-E transitions only) and initial earnings category (for E-to-U transitions only). Since this combination of characteristics is more coarse than the actual set of observed characteristics used to define transition groups in the LEHD, I use the (sometimes spurious) U-to-E, E-to-U, and U-to-U transitions in the LEHD only to distribute the true (up to ACS sampling error) counts across the disaggregated transition groups I use. I supplement the ACS data with BLS data on national unemployment rates by age group to ensure that the scale of the labor force is consistent with standard measures. Appendix A5 provides further detail about this imputation procedure.35

The second problem is that excluding job-to-job transitions into or from states outside of the 19 state sample will cause the counterfactual simulations to overstate the geographic concentration of demand shock incidence, since workers from the remaining states are effectively excluded from competing for the new positions. Because accurate forecasts of shock incidence is likely to be particularly sensitive to observing flows of workers to and from states adjacent to or near the state receiving the labor demand shock, for the purpose of the counterfactual simulations below I only choose target census tracts from a subset of 10 preceding and following quarter as well.

35Note that this imputation procedure also solves an additional problem: by only including spells of nonemployment that are bookended by spells of observed employment, my LEHD sample in isolation is likely to severely undercount U-to-U and E-to-U transitions in the last few years of the sample, since true unemployed workers that remain in the labor force have not had to time to find jobs that end their unemployment spells. However, under the imputation procedure, the ACS and BLS data set the scale of U-to-E, E-to-U, and U-to-U transitions at a slightly aggregated level rather than relying on the biased LEHD counts.
states in the west/southwest/great plains area where I have near complete coverage, so that almost all nearby states are observed (though the full 19 state sample is always used to represent the “national labor market” in all simulations to minimize the remaining bias). Note though, that despite this likely overstatement of the within-state share of shock incidence, I nonetheless find below that a large share of the incidence of the shocks I consider accrues to initially out-of-state workers.

4.3 Assigning Job Matches to Types and Job Transitions/Retentions to Groups

For each pair of years \((y - 1, y)\) I assign each job transition/retention observation to an origin type \(o^{y-1}(i, j)\), a destination type \(d^y(k)\), and a transition group \(g^{y-1:y}(i, j, k)\) (time superscripts will henceforth be dropped except where necessary). Specifically, a worker \(i\) with primary year \(y - 1\) position \(j\) is assigned to an origin type \(o(i, j)\) based on the combination of the location of the establishment \((m(j))\) associated with position \(j\) (discussed further in Section 5.2 below) and the earnings quartile while at primary position \(j\). If worker \(i\) was not employed in year \(y - 1\), the worker is assigned to one of two unemployed types, differentiated by age \( (< 25 \text{ or } \geq 25)\) to distinguish new entrants/recent graduates from workers with meaningful work experience, since employers might treat new and experienced unemployed workers as quite imperfect substitutes for each other. The same worker \(i\) with year \(y\) primary position \(k\) is assigned to a destination type \(d(k)\) based on the combination of establishment \(m(k)\)’s geographic location, \(m(k)\)’s size quartile (based the establishment employment distribution), its quartile of average worker earnings (again using an establishment-level distribution), and its industry supersector. These characteristics were chosen because they are consistently observable and likely to be important determinants of productivity complementarities, geographic recruiting and search costs, and the other components of the relative job match surplus a destination position will have with workers of alternative skill levels from different locations. The pair of primary jobs for worker \(i\), \((i, j, k)\), is assigned to a group \(g(i, j, k) \equiv g(o(i, j), d(k), z(i, j, k))\) based on the origin job type \(o(i, j)\), the destination job type \(d(k)\), and an indicator \(z(i, j, k)\) for whether the establishments of positions \(j\) and \(k\) are the same \((m(j) = m(k))\).

\[36\]Earnings quartile cutoffs are defined relative to the distribution of primary job annual earnings for workers in the state-year combination associated with the observation, and are based on prorating earnings in full quarters only to ensure that the quartile better captures a worker’s skill rather than the share of the year in which he/she was working. One drawback of the LEHD database is that a worker’s location must be imputed for multi-establishment firms. However, the Census Bureau’s unit-to-worker imputation procedure assigns an establishment to a worker with a probability that decreases in the distance between the worker’s residence and that establishment. Consequently, cases of significant measurement error in true location are unlikely to occur, since most mistakes will misattribute the worker’s job to an establishment within the same tract or perhaps a nearby tract.
### 4.4 Summary Statistics

Figure 1 (Table A2, Col. 1) displays the raw distribution across distance bins of the distance between the geographic location of origin and destination establishments for workers who changed dominant jobs ($m(j) \neq m(k)$) between 2010 and 2011. 3.5% of job movers moved to a different job within the same census tract, while another 7.6%, 7.2%, and 13.8% percent moved to a job one tract, two tracts, or 3+ tracts away within the same PUMA. 60% found a job in another PUMA within the same state, while 8% changed states. The sizable share of workers moving to new jobs located very close to their previous jobs is prima facie evidence that either search/moving costs are large or preferences for particular locations are strong, so that conditions in a worker’s local labor market may still hold outsized importance for their job-related welfare.

Table A3 Panel A, Row 1 shows that about 21.3 percent of year-to-year pairs of dominant jobs in the sample involve moves to new jobs (either job-to-job transitions or unemployment-to-employment transitions). A full 70.3% of workers keep the same dominant job from one year to the next, while 4.2% and 4.3% of year-to-year pairs involve unemployment-to-unemployment and employment-to-unemployment transitions, respectively. Collectively, the 2010-2011 sample used to estimate the set $\{\Theta_{D-in-D}\}$ used for the counterfactual simulations reflects 23 million year-to-year pairs of dominant jobs.

Examining other rows of Panel A, we see that about 85% of workers younger than 25 who were unemployed in 2010 found jobs in 2011, while only 53.1% of older unemployed workers transitioned to a job in 2011, highlighting the importance of treating these two groups of unemployed workers asymmetrically. We also see that employed workers in the lowest earnings category in 2010 were far less likely to stay at their job (66.7%) and far more likely to transition to unemployment (9.7%) or another job (23.6%) than those with higher paying jobs. By contrast, 87.3% of those in the highest earnings quartile in 2010 stayed at their job, while only 10.2% transitioned to new jobs. However, conditional on changing jobs, the highest earning workers were most likely to leave their original PUMA and most likely to change states, suggesting that the geographic scope of labor markets may differ across skill categories. These differences provide motivation for including nonemployment status/earnings quartiles as characteristics used to define origin worker types.

Panel B of Table A3 shows that higher paying firms are considerably more likely to retain their workers, but are also slightly more likely to hire more distant workers when they do fill a vacant position: for establishments in the highest average pay quartile, 9.6% of new workers were working out of state the previous year and 29.2% were working in the same PUMA, while for establishments in the lowest average pay quartile, 7.5% were working out of state and 32.3% were working in the same PUMA. The largest quartile of establishments (as measured by employment) are also more likely to retain their workers, but are the least likely to hire from out of state (4.3%). Establishments in the smallest size
quartile hired a whopping 27.2% of new workers from out of state, and only 28.2% from the same PUMA, suggesting, perhaps counterintuitively, that the smallest establishments seem to operate in the most integrated labor markets. While these statistics motivate the choices of types and the need to consider labor demand shocks featuring different establishment compositions, they do not condition on any other firm, location, or worker characteristics. Comparing incidence across counterfactual shocks that only differ along one dimension of establishment characteristics will be far more informative about the relative scope of labor markets across different types of workers and establishments.

5 Estimation

5.1 Defining the Local Labor Demand Shocks

I consider three categories of local labor demand shocks: stimulus packages, establishment relocations, and natural disasters. Each stimulus package shock consists of 500 jobs that are added to the destination-year stock of jobs to be filled in a chosen census tract, combined with the removal of 500 unemployment “positions”. Given that census tracts have on average around 5,000 jobs, this represents about a 10% increase in labor demand for the average tract. For each chosen tract, I simulate 35 different stimulus packages, each differing in the particular kind of establishment whose demand increases by 500 jobs based on combinations of the remaining non-location establishment attributes that define a destination type: quartiles of establishment size and average pay along with industry supersector. The full list of the establishment compositions of the simulated shocks is displayed in Table A1. The compositions were chosen to highlight the heterogeneity in incidence across different industry/size/avg. pay cells. All but three of the stimulus packages place no restrictions on which workers the new establishments may consider when filling their new positions. For the remaining three, I restrict the new establishments to fill the new positions using only workers from the surrounding PUMA, so as to reflect the kinds of stipulations sometimes included in economic development contracts between cities and incoming firms. Comparing these “restricted” specifications to their unrestricted counterparts illustrates the value to cities or states of including such provisions in agreements.

The establishment relocation shocks identical in structure to the stimulus packages, except that instead of subtracting 500 positions from the nonemployment destination type, the 500 positions are instead subtracted from a sending destination type associated with the same establishment characteristics as the type receiving the jobs but located at least two states away. This ensures that locations near to the “winning” site are minimally affected by the lost employment at the “losing” site. Unlike the stimulus package shocks,

\footnote{For example, under the Empowerment Zone system, firms only received wage subsidies for workers residing in the local area (Busso et al. (2013)).}
which increase nationwide labor demand, the relocation packages merely redistribute existing demand across locations. If there were no spatial search frictions and workers could move costlessly, then such shocks would merely cause the moving establishment’s workers to follow the establishment, no additional reallocation would be necessary, and the “shock” would have zero incidence for workers in all initial skill groups and locations. By contrast, if mobility costs were so high as to eliminate mobility between small local labor markets, the impact on local workers of a stimulus package and a relocation with the same establishment composition would be identical. Because the relocation packages ended up having quite similar incidence to their stimulus counterparts, I restrict attention to three representative specifications.

Finally, I also consider “natural disaster” shocks in which a targeted census tract loses a random 25%, a random 50%, or all 100% of its jobs in the destination year, with the number of lost jobs being added as “positions” to the nonemployment destination type. These simulations provide an opportunity to examine whether the skill incidence of negative shocks is symmetric to positive shocks, as well as to consider the degree to which higher skilled workers initially working in the targeted tract are able to capture a disproportionate share of the remaining local jobs when only a fraction of such jobs are eliminated. These disaster simulations are also included to illustrate how the two-sided matching model could be customized to handle any particular disaster scenario, including disasters such as hurricanes that hit a number of contiguous tracts simultaneously (and perhaps with differential force).

5.2 Collapsing the Type Space for Distant Geographic Areas

While the census tract is the smallest unit of geographic aggregation I use to define geographic location, given that the 19 state sample features nearly 28,000 census tracts, and transition groups \( g \) are defined by several other worker and establishment characteristics in addition to origin and destination locations, treating all census tracts as separate locations will generate trillions of transition groups. This in turn will cause the empirical group distribution \( \hat{P}(g) \) to be too sparsely populated to allow precise estimates of the set of relative group surpluses \( \Theta^{D-in-D} \).

Thus, since I am particularly interested in the incidence of demand shocks of alternative compositions across locations relatively near to the site of the shock, I combine initial groups that are defined by the same worker and establishment characteristics and are geographically proximate to each other but far from the site of the shock. Specifically, given the focus on carefully measuring local incidence, I do not combine any origin or destination types featuring a location within a five tract pathlength of the targeted tract for a given simulation. However, I combine into a single origin (or destination) type any origin (or destination) types that share the same worker (or establishment) characteristics and feature tracts that are in the same state as the targeted tract and in the same public-use microdata area (PUMA) as
each other (but not the same PUMA as the targeted tract). Thus, outside a 5-tract circle surrounding the targeted tract, the geographic locations of types are defined by PUMAs rather than tracts. Furthermore, for types featuring tracts outside the targeted state, I combine types featuring the same worker (or establishment) characteristics whose tracts are in the same state. Thus, outside of the targeted state, the geographic locations of types are defined by states rather than PUMAs.

Coarsening the type space for distant geographic locations dramatically decreases the overall number of groups and the severity of the sparse matrix problem. In particular, while many job-to-job transitions are between nearby tracts, there are very few transitions between any chosen pair of tracts in distant states, so that relative surplus parameters for transition groups featuring tracts in different states would never be well-identified without such coarsening. Note that this approach still incorporates all of the transitions and all of the locations in the 19 state sample into each simulation, so that each local labor market is still nested within a single national labor market.

Even after the type space has been coarsened, there are still relatively few observed transitions per remaining transition group \(g\), particularly for groups local to the shock. Thus, I further address the sparsity of the empirical group-level distribution \(\hat{P}(g)\) by smoothing this distribution prior to estimation by making each element’s value a kernel-density weighted average of groups featuring “similar” worker and position characteristics. Because excessive smoothing across other transition groups erodes the signal contained in the data about the degree of heterogeneity in the relative surpluses from job transitions featuring different combinations of worker characteristics and establishment characteristics and locations, I take particular care when devising the smoothing strategy.

Full detail about this customized smoothing procedure is provided in Appendix Section A6. However, the procedure is based on the intuition that the geographic location of a destination establishment is likely to be critical in determining the origin locations whose associated worker transitions generate the most surplus (i.e. least moving/search cost), while the combination of non-location characteristics (establishment size, average pay, and industry) is likely to be more important than location in determining the skill category of worker (proxied by initial earnings quartile) that generates the most surplus.

Note that the type aggregation and smoothing procedures imply that the origin and destination type space will be different for each simulation that involves a different targeted tract. Furthermore, disclosure restrictions imposed by the FSRDC system prevent the release of any results that are specific to a particular substate geographic location. Thus, while each simulation is performed with a particular tract level target, I only report averages of incidence measures across 500 simulations for each shock type, where each simulation features a different randomly chosen target census tract from the 10 state southwest/west/great plains subsample.\footnote{I impose that a census tract is only eligible to be a target tract in the simulations if it features at least 100}
is used for each of the alternative simulations, so as to facilitate fair comparisons between the alternative stimulus packages, establishment relocations, and natural disasters.

After the simulations have been run, the space of transition groups is again redefined in order to average simulation results across alternative targeted census tracts. This time, origin and destination type locations are replaced with bins capturing distance to the targeted census tract, and I report estimates of incidence for various distance rings around the site of the shock. Note, though, that during the simulations themselves the spatial links between adjacent and nearby tracts are not restricted to follow a particular parametric function of distance between locations. Thus, to this point no prior assumption about the role of distance has been imposed during estimation beyond the initial aggregation of distant tracts to PUMAs and states described above.

5.3 Standard Errors

While this draft uses a 50% subsample of all pairs of dominant jobs in adjacent years in the population of 19 states I consider, the next draft will use the full 100%. Thus, it is not obvious how to define the relevant population for the purposes of inference. Furthermore, since I estimate millions of surplus parameters \( \theta_g \in \Theta \), and each counterfactual incidence statistic depends on the full set \( \Theta \), I do not report estimates of any single parameter. Instead, any standard errors that might be reported ought to provide information about the precision of the incidence forecasts. One source of sampling error that can be quantified relates to the particular random draw of 500 census tracts used as the targets for the local labor demand shocks. In the first few results tables discussed below, I provide standard errors that reflect the sampling error generated by the fact that the estimates average over only one possible draw of 500 census tracts, while the object of interest is the average among all census tracts in the sample (around 28,000). These standard errors prove to be tiny, suggesting that there is little value to running additional simulations featuring additional target tracts for the purpose of better estimating expected incidence for a randomly drawn tract. As a result, subsequent tables do not report these standard errors at all. Note, though, that while these standard errors capture the variation in estimated incidence across alternative target locations, they do not distinguish whether this variation is driven by true sampling error stemming from a limited number of observed transitions per \( \theta_g \) parameter estimated versus systematic differences in worker and establishment productivities, moving costs, and non-wage utilities across locations driven by past sorting.

39 When defining types and when presenting results I generally focus on distance bins defined by tract, PUMA, and state pathlengths rather than miles, since the number of potential workers contained within circles defined by the same tract or PUMA pathlength is likely to be more consistent across urban and rural areas featuring very different densities than circles whose radius is defined by a number of miles.
6 Results

6.1 Stimulus Packages

6.1.1 Incidence by Distance to Focal Tract

Before comparing the impacts of stimulus packages featuring different establishment compositions, I focus first on characterizing the geographic scope of labor markets for a “typical” local stimulus. I do this by averaging the predicted change in worker-position allocations produced by the stimulus packages across all 32 stimuli simulated, effectively integrating over the joint distribution of establishment industries, sizes, and average pay levels. While I focus attention on graphical representations of the results contained in a set of figures, most figures have an accompanying Appendix table (listed in parentheses) that contains the particular values that were plotted and are cited in the text.

Figure 3a (Table A4, Col. 1) illustrates the mean probability of receiving one of the 500 new stimulus jobs for randomly chosen individuals initially working at different distances from the census tract receiving the stimulus. The figure reveals the sense in which U.S. labor markets are still quite local: the probability of obtaining one of the new jobs for a worker initially working within the target tract (.015) is three times higher than for a worker working in an adjacent tract (.005), over 7 times higher than for a worker working 2 tracts away (.002), and 15 times higher than for a worker working 3 or more tracts away within the same PUMA (.001). Furthermore, additional distance from the focal tract continues to matter at greater distances: the probability of obtaining one of the stimulus jobs for a local (target tract) worker is 27 times higher than for a worker in an adjacent PUMA, 55 and 283 times higher than for a worker two PUMAs away or 3 or more PUMAs away within the same state, respectively, and 1,087 and 9,375 times higher than for a random worker one state away or 2 or more states away, respectively.

However, the vast differences in $P(\text{new job} \mid \text{distance from target})$ present a very misleading guide to the overall incidence of new jobs across geographic locations. This is because the target tract initially employs an extremely small fraction of the population in the 19 state sample who are defined by the model to be at risk of obtaining the stimulus jobs. Figure 2 shows the share of workers in the simulation samples that are working in each distance bin relative to the targeted census tract prior to the stimulus. Only 0.0045% of the workforce is composed of workers initially working in the target tract. As expected, the shares get larger quickly as one considers distance bins defined by concentric circles with much larger radii: 0.026%, 0.059%, 0.3% of the workforce initially work 1, 2, or 3+ tracts away from the target tract within the same PUMA, while 0.4%, 1.1% and 16.2% initially work 1, 2, or 3+ PUMAs away, respectively, and 13.5% and 68.4% initially work 1-2 states or 3+ states away.

Consequently, if one swaps the terms in the conditional probability and calculates the
share of stimulus jobs obtained by workers initially working or nonemployed in each of the distance bins listed above, \( P(\text{distance from target} \mid \text{new job}) \), one obtains a very different impression of incidence. Figure 3b (Table A4, Col. 2) displays the mean share of new jobs by distance bin across the 32 simulated stimulus packages. 3.1% of new jobs go to workers initially in the target tract, another 16.2% are obtained by other workers in the PUMA, 60% are obtained by workers in different PUMAs within the state, and 10.6% are obtained by out of state workers. So a very large share of the new jobs are likely to be obtained by workers far outside the local jurisdiction that is hosting the stimulus (and is likely lobbying for its local placement).

One could likely obtain similar forecasts of the shares of workers by distance bin who would obtain jobs at a new establishment simply by looking at the distance composition of workers who obtained jobs from actual plant openings in the past. As emphasized in the introduction, though, the probabilities of obtaining the particular new jobs created by the stimulus package may not be particularly informative about the true incidence of the shock. This is because many of the workers who obtain the new jobs would have obtained other similarly paying jobs in the absence of the stimulus, and nearby workers may now obtain the jobs these workers would have accepted or retained, and so on, creating ripple effects through vacancy chains that determine the true employment and welfare incidence. This is where the use of a flexible equilibrium model is particularly valuable.

Figure 4a (Table A4, Col. 3) is analogous to Figure 3a (Table A4, Col. 1), except that instead of the probability of obtaining a particular stimulus job, it captures the change in the probability of any employment (or equivalently, the change in the probability of unemployment) due to the stimulus, relative to a no-stimulus counterfactual, for randomly chosen workers initially working at different distances from the target census tract of the stimulus.

The figure demonstrates that the change in employment probability is still quite locally concentrated, though less so than the probability of obtaining a stimulus job. Workers initially employed (or nonemployed) in the target tract are 0.3% more likely to be employed at the end of the year than in the absence of the stimulus. This is 3.2, 5.0, and 7.8 times greater than the corresponding changes in employment probabilities for workers 1, 2, or 3+ tracts away (within the same PUMA), 8, 17, and 49 times greater than for workers 1, 2, or 3+ PUMAs away (within the same state), and 150 and 448 times greater than for workers one state and 2+ states away, respectively. In particular, the odds of changes in employment status for workers 2+ states away relative to workers in the local tract are 21 times higher than they were for the probability of obtaining a stimulus job.

The broader geographic incidence for general employment status is reflected in Figure 4b (Table A4, Col. 4), the analogue to Figure 3b (Table A4, Col. 2), which displays the share of the aggregate 500 job increase in employment that accrue to workers initially employed in each distance bin relative to the target tract. Only 0.5% of the net employment change
redounds to workers initially employed in the target tract, with 7.6% of the additional employment going to workers in other tracts within the PUMA, 57.4% to workers in other PUMAs within the target state, and 34.5% going to workers initially employed out of state.

The simulation procedure also generates counterfactual changes in worker mean job-related utility necessary to clear the market for each origin job type following the various stimuli. Recall, though, that generating the market clearing allocation only requires computing standardized utility premia $\frac{C_o - C_o'}{\sigma}$. As discussed in section 3.4 and Appendix A3, I exploit the existence of a longer panel of years to generate estimates of $\sigma$ that, given the assumption of a money-metric utility function, allow utility premia to be scaled in dollars of annual earnings. However, the assumptions that underlie the estimate of $\sigma$ are stronger than for the relative joint surplus values (and are extremely unlikely to hold exactly). Thus, while proportional earnings changes for different skill and location categories ought to be reasonably well-identified, the estimated dollar value of predicted welfare gains should be treated cautiously. Furthermore, since only relative utility changes are identified, I normalize the estimated utility impact to be 0 for the origin worker type estimated to experience the smallest impact (which varies by the composition of the stimulus package, but is generally initially low paid workers in some distant state), so that all estimated utility changes are relative to this origin type.

Figure 5a (Table A4, Col. 5) provides the average utility impact (scaled in annual earnings equivalents) for random workers initially employed (or unemployed) at different distances from the target tract for the “typical” stimulus package (again averaging over all 32 simulated stimuli). Workers initially working in the focal tract receive an estimated $1,045 increase (in 2011 dollars) in money metric utility from the typical stimulus package (relative to the least affected origin worker type), while workers initially working 1, 2, and 3 or more tracts away receive expected utility gains of $395, $278, and $164 respectively. Workers initially working 1, 2, and 3+ PUMAs away within the state receive the utility equivalent of $164, $143, and $109 in annual earnings gains, while workers one and 2+ states from the site of the shock receive average gains of $89 and $85. Figure 5b (Table A4, Col. 6) plots the share of total utility gains (relative to distant workers) that accrue to workers in each distance bin. Only 0.5% of total worker welfare gains accrue to workers within the focal tract, with another 6.5% accruing to other workers originally within the PUMA associated with the focal tract. 51.5% of the gains accrue to workers outside the targeted PUMA but within the same state, while 41.6% accrue to workers initially employed out of state. Thus, examining incidence from the perspective of welfare gains rather than employment gains suggests an even more geographically integrated labor market.

Table A5 provides the expected employment and welfare gains and shares of total employment and welfare gains accruing to workers in each distance bin when distance bins are constructed based on miles from the focal tract rather than based on tract, PUMA, or state pathlength. The story is essentially the same: only 10.3% of net employment gains and
only 11.6% of welfare gains accrue to workers initially working (or most recently working) within 10 miles of the target tract even though 32.3% of stimulus jobs are filled by such workers. 56.5% of employment gains and 53.3% of welfare gains accrue to workers initially working more than 250 miles from the site of the shock.

6.1.2 Heterogeneity in Incidence Across Establishment and Worker Characteristics

Figure 6a (Table A6, Col. 2-9) illustrates how the change in employment probability for random workers in different distance bins varies across stimulus packages featuring new positions in different industry supersectors, while Table A7, Col. 2-9 shows how the share of the net employment change accruing to workers in each distance bin varies by the industry composition of the shock. The figure shows that the employment incidence across distance bins is shockingly similar across shocks featuring different industry supersector composition. Some small differences do exist: 8.8% and 66.5% of the change in net employment goes to workers within PUMA and within state, respectively, in stimuli featuring positions in the other services sector (which includes repair and maintenance, personal and laundry services, and religious/civic organizations) sector, which is the sector featuring the most geographically concentrated employment incidence. By contrast, the corresponding figures are 7.4% and 65.1% for stimuli featuring positions in the retail/wholesale trade supersector (among the least geographically concentrated). Analogous results for the expected change in worker utility show somewhat more industry-level heterogeneity in incidence, particularly for the most local workers (Table available upon request): stimuli featuring other services jobs vs. those featuring retail/wholesale trade jobs differ in expected welfare gain for those initially in the focal tract by nearly $400 in money-metric utility (1344 to 955).

Figures 7a and 7b (Table A8) capture the corresponding heterogeneity in geographic incidence across stimulus simulations featuring positions at establishments from different establishment size quartile/establishment average pay quartile combinations (averaging over industry supersectors). On average, stimuli featuring positions at large establishments (4th quartile) with low average pay (2nd quartile) generate the most local incidence (0.6%, 8.5%, 59.7%, and 31.2% of net employment gains within tract, among other tracts within PUMA, among other PUMAs within state, and among other states, respectively), while small establishments (1st quartile) with high average pay (4th quartile) generate the least local incidence (0.5%, 7.8%, 54.4%, 37.9%). These differences are modest but nontrivial. Table A9 shows the corresponding estimates for the expected change in worker utility. For this outcome, stimuli featuring jobs at small, low paying establishments generate the greatest local incidence, though differences are again modest.

40I chose the 2nd quartile of firm average pay to represent “low paying establishments” rather than the 1st quartile so that the stimulus packages would be considered desirable for the receiving tract (most local development initiatives do not seek to increase the number of minimum wage jobs).
Another feature of the model is the ability to capture heterogeneity in shock incidence across workers in different skill classes, as proxied by initial employment status and earnings quartile. Figure 8a (Table A10, Col. 1) captures the share of the 500 job net employment gain that is enjoyed by workers whose initial earnings fall in each quartile of the national distribution, as well as workers who were unemployed in the year prior to the simulated stimulus shock. 41.1% of the employment gains accrue to those initially unemployed (8.9% to workers younger than 25, 32.2% to older workers), while 26.1%, 15.4%, 9.6%, and 7.6% accrue to those at the 1st through 4th quartiles of the initial earnings distribution, respectively. The smaller values for initially high earning workers reflect the fact that such workers were less likely to transition to unemployment in the absence of the shock. Figure 8b (Table A10, Col. 6) displays the share of the total worker welfare gain enjoyed by workers in each initial earnings quartile (and initially unemployed workers). For the typical shock, 10.3% of utility gains accrue to initially unemployed workers, while the share accruing to each earnings quartile increases in the level of initial earnings: 20.1%, 21.4%, 22.3%, and 26.0% for quartiles 1-4, respectively. These results suggest that existing high paid workers receive a disproportionate share of the welfare gains from a typical shock.

When averaged across all distance categories, there is very little heterogeneity in skill incidence across stimuli featuring different industry supersectors: all supersectors feature shares of utility (employment) gains accruing to each nonemployment/initial earnings category that fall within 1% (4%) of the overall average across all stimuli (table available upon request). Surprisingly, changing the firm size/pay categories associated with the stimulus jobs also produces quite limited heterogeneity in incidence across initial earnings categories at the national level. Stimuli featuring low paying firms only generate 1-2% higher shares of employment and welfare gains for initially low paid workers relative to those featuring high paying firms (Table A10). By contrast, only ~24% of stimulus jobs at high paying firms are predicted to be obtained by initially unemployed workers versus ~30% for low paying firms, and ~34% versus 7% of jobs from high versus low paying firms go to workers initially in the highest earnings quintile, suggesting that the skill incidence of the actual stimulus jobs understates the degree to which employment gains “trickle down” to initially nonemployed workers from labor demand shocks featuring a bias toward high skilled workers.

One can also examine the degree to which the geographic scope of labor markets depends on the skill level. Figure 9a (Table A11) examines the change in employment probability for a randomly chosen worker whose initial job (or nonemployment) places him/her in a particular earnings quartile/distance bin combination. We see that older unemployed workers who most recently worked in the focal tract enjoy a sizable decrease in nonemployment rate of 1.6 percentage points, while the unemployment rate decrease is only 0.4% and 0.3% for workers most recently employed one or two census tracts away, indicating that the employment gains for existing unemployed workers are particularly local. That said, employment gains decline with distance in a relatively similar fashion for all initial earnings quartiles.
Existing older unemployed workers in the target district enjoy 0.2% of the total employment gains despite only constituting 0.0003% of the population of potential workers.

Figure 9b (Table A12) displays the analogue to Figure 9a (Table A11) for welfare changes. Experienced unemployed workers most recently employed in the targeted census tract enjoy a change in utility equivalent to $1165 in annual earnings from a typical shock, while newer entrants to the labor market experience much smaller gains ($620), partly because they were more likely to find jobs in the absence of the stimulus. Among the initially employed, welfare changes rise monotonically from $999 for the 1st initial earnings quartile to $1242 for the 4th. Welfare gains decrease more quickly with distance from the focal tract for the higher income groups, however, creating rapid convergence in welfare gains across different earnings quartiles with distance from the focal tract. Thus, the results from Figure 8b (Table A10, Col. 6) that averaged across all distance categories when examining overall welfare incidence across skill levels obscured the much larger differences in welfare incidence between that occurs among differently skilled workers local to the shock.

Moreover, aggregating across all distance categories also obscures substantial heterogeneity across shocks featuring different establishment composition in the welfare incidence among skill categories for workers in the focal tract. Figure 10a, (Table A13, Col 2-9) shows the money metric utility gains for only local workers by initial employment status/earnings quartile of the worker and industry supersector of the stimulus. Typical leisure/hospitality and other services stimuli yield welfare gains for existing older unemployed workers equivalent to $1,374 and $1,413 in annual earnings, compared to $952 and $1000 for stimuli featuring new jobs in the information or state/local government sectors. Workers in the highest (4th) initial earnings quartile reap expected utility gains of only $1025 from stimuli featuring jobs in the education/health supersector, while other services stimuli generate $1640 for such workers (as noted above, stimuli in the other services supersector feature particularly local incidence for all skill levels). The rankings of utility gains for local workers across industries also differ strikingly across different unemployment/initial earnings categories, with construction-centered stimuli offering the lowest payoff for young nonemployed workers, retail/trade and leisure/hospitality the lowest for initially low-paid workers, and education/health offering the lowest payoff for initially high paid workers.

Figure 10b (Table A14) shows the corresponding expected utility gains for workers in the focal tract by establishment size/establishment pay quartile combinations instead of industry. Not surprisingly, high-paying firms (regardless of size) generate much larger gains for initially high paid workers and smaller gains for both initially low-paid and nonemployed workers. Stimuli featuring positions at small, high paying firms generate the least payoff for initially nonemployed workers: $413 and $1015 for young unemployed and older unemployed workers, respectively, while generating a substantial $1728 for 4th quartile workers. Large, low paying firms show the opposite pattern, with payoffs of $802, $1297, and $872 for the same three groups. In general, smaller firms seem to generate larger gains for previously low-
paid but employed workers while larger firms generate more gains for initially unemployed workers. Thus, assuming similar impacts on rent price and feedback effects through the product market, it appears that the skill requirements of the positions being created matters a lot for incidence among skill classes for local workers, but much less for workers farther away.

In addition, the substantial heterogeneity in local skill incidence across industries and establishment size/pay quartile combinations still misses further heterogeneity operating at the three-dimensional supersector/size/pay cell level. Figure 11 (Tables available upon request) plots the forecasted welfare gains by initial employment status/earnings quartile among workers initially in the focal tract for all 32 stimulus shock compositions that were simulated. There is a huge range of predicted gains. Welfare gains for younger initially unemployed workers range from $284 (small, high paying government positions) to $1052 (large, low paying leisure/hospitality positions), while their older counterparts exhibit a range from $674 (also small, high paying government positions to $1750 (small, low paying other services). For 1st earnings quartile workers, they range from $718 (large, high paying information positions) to $1703 (small, low paying other services). For 4th quartile workers, they range from $696 (large, low paying leisure/hospitality positions) to $2376 (small, high paying other services positions). For small precinct councilors whose concern is primarily very local incidence, these represent massive differences in the scale and skill intensity of the earnings incidence.

Besides the characteristics of the establishment(s) that constitute the stimulus, another important source of heterogeneity in geographic incidence is the residential population density and employment density of the area receiving the shock. Figure 12 (Tables A16 and A15) provides the average employment and welfare incidence among the least dense (most rural) 100 and most dense (most urban) 100 of the 500 tracts receiving simulated shocks, as well as among the 100 tracts with smallest and largest pre-shock employment levels (as measured by employees among establishments located in the target tract). The 100 least dense tracts tend to have residential population below 500, while the 100 most dense tracts have residential populations greater than 8,000. Similarly, the 100 tracts with the smallest pre-shock employment levels have fewer than 250 workers originally, while those with the largest have more than 8,000.

Both employment and welfare gains are more geographically concentrated for tracts featuring lower population density. The average share of employment gains enjoyed by workers initially within the targeted PUMA is 9.5% among the most rural 100 tracts versus 5.6% for the 100 most urban tracts (and 7.1% among all 500 tracts simulated). The measures of welfare incidence tell essentially the same story, with the expected money-metric utility gain for workers within the focal tract, 1 tract away, 2 tracts away, and 3+ tracts within the PUMA all considerably larger for the most rural focal tracts relative to the most urban focal tracts ($1724 vs. $799, $639 vs. $175, $431 vs. $155, and $250 vs. $120, respectively).
The differences in expected welfare gains among nearby workers are even larger for tracts featuring a small number of initial workers relative to those featuring many (e.g. $1790 vs. $434 for those already working in the focal tract). Interestingly, though, despite their disproportionate gains, there are so few workers already working in the 100 smallest tracts by initial employment relative to the largest 100 tracts that the share of all welfare and employment gains accruing to workers within the focal tract is greater in the latter set of tracts (0.3% versus 0.7% for both outcomes).

Figure 13 (Table A17) illustrates the impact on stimulus incidence of requiring the incoming establishments to fill their positions using only workers originally working (or most recently working) in the same PUMA as the tract receiving the stimulus. The first two columns of the table reveal that the unemployment probability falls by 1.4% instead of 0.3% for a randomly chosen worker in the target tract when hiring is restricted to occur within PUMA, while the decrease in unemployment probability is 2-3 times as large in the restricted vs. unrestricted version of the stimulus for workers initially working outside the focal tract but within the PUMA. Overall, the within-PUMA share of net employment gains increases from 9.5% to 23.5% when hiring is restricted to those already in the chosen PUMA (col. 3-4).

Restricting hiring also increases the expected money metric utility gains by seven fold ($995 to $6938) for a randomly chosen worker already or most recently working in the focal tract, with 2-4 fold increases in utility gains for other workers initially in the PUMA, depending on initial distance to the target tract. The share of utility gains accruing to the targeted PUMA increases from 7.2% to 17.5%. Thus, it seems that local development initiatives such as empowerment zones that add stipulations restricting hiring decisions or wage subsidies to only the local workers are likely to succeed in generating a much more locally concentrated labor market incidence. Note that this is despite the fact that any additional downstream hiring caused by initially employed workers vacating jobs to move to the 500 new positions remains unrestricted in these counterfactuals.

As noted in the estimation section, “plant relocation” shocks were also simulated that feature the same compositions of 500 new jobs as the stimulus packages but removed the jobs from a distant state rather than from the stock of “nonemployment” positions. The average welfare gain/loss and increase/decrease in probability of employment are presented in Table A18 for two particular relocations involving 500 positions at small, high paying information establishments and large, low paying manufacturing establishments, respectively. Since the “losing” locations are so far away from the “winning” tracts, and as shown above labor markets are still quite local and (to a lesser extent) regional, the employment incidence of such relocation shocks is virtually identical to their stimulus counterparts for locations within the winning state, and the small welfare differences between stimulus and relocation shocks are due purely to the change in normalization.\(^{41}\)

\(^{41}\)For stimulus packages, the origin type with smallest predicted welfare gain is normalized to 0, while for
6.2 Natural Disasters

Recall that the “natural disaster” simulations remove at random 25%, 50%, or 100% of the destination jobs in the focal tract. Averaging over initial earnings categories, Figure 14a (Table A19, col. 1-4) displays the increase in the probability of unemployment for randomly chosen workers within different distance bins from the focal tract for each disaster intensity. Workers initially working (or unemployed) in the focal tract experience increases in the probability of unemployment in the destination year of 2.9%, 7.3%, and 19.2% from the 25%, 50%, and 100% disasters, respectively. The share of new unemployment that falls upon workers initially in the local tract increases from 12.6% when 25% of local jobs disappear to 20.2% when 100% of local jobs disappear, suggesting that the employment incidence becomes increasingly geographically concentrated the more intense the local disaster (even when the disaster itself is in each case still contained within the same census tract) (Table A19, col. 5-8). As with the stimulus shocks, in one sense these results suggest that labor markets remain very local: if mobility among labor markets were truly frictionless (and workers/positions were homogeneous), so that the predicted employment incidence fell equally across all workers, then the expected share of lost jobs borne by local workers would be their share of the total workforce: 0.0045%. Thus, local workers experience a change in nonemployment probability that is over 4,000 times larger than it would be in a frictionless, homogeneous world.

Note also that the employment incidence of the simulated disaster is more locally focused than for the stimulus packages, as measured by the local share of the total employment change. In the case of stimulus packages, most of the local workers would have been working (somewhere) in the absence of the shock (or are long-term unemployed workers that would produce little joint surplus from employment, due to either preferences or low productivity), so that there was an effective limit to how local the employment incidence could be, thus forcing much of the net employment gain to be distributed across more distant locations. Analogously, since the unemployment rate is well below 50% in most locations before the simulated shock, very few would have become unemployed in the absence of the disaster, creating scope for a large effect. Indeed, the shocks produce a greater local labor surplus than existed nearly anywhere in the data; since most positions retain their existing workers (appearing in the model as a large increase in joint surplus from a match when the individual is an existing employee), it is very difficult for all of the local workers to find jobs. Thus, the model estimates reveal a natural asymmetry in the geographic scope of employment incidence between positive and negative local demand shocks.

Figure 14c (Table A19, col. 9-12) shows the average utility losses by distance bin for each disaster intensity. Expected utility losses (scaled as equivalent annual earnings losses) are severe for workers initially in the focal tract: $-5,622, $-10,474, and $-17,028 for disasters
featuring 25%, 50%, and 100% local job loss, respectively (relative to the least affected origin worker type nationally). Because losing 100% of jobs is fairly far outside the support of the shocks observed in the data, and employment expansions by establishments taking advantage of a local labor surplus have been shut down in the model, these massive welfare losses should be considered somewhat skeptically. The welfare losses fall dramatically to $-235, $-361, and $-543 for workers in an adjacent tract, and then decrease slowly in magnitude to $-115, $-135, and $-135 for those initially working outside the targeted state. However, because within-tract workers are such a small share of the working population, the shares of aggregate worker welfare losses accounted for by the losses of workers initially employed in the focal tract are still only 10%, 9.7%, and 8.3% for the three disaster intensities, respectively (Figure 14d, Table A19, col. 13-16). While this local share is 20 times larger than the local share of welfare gains for the stimulus packages, it is nonetheless fairly small. As before, over 80% of the earnings incidence is predicted to fall on out-of PUMA workers and around 40% on out-of-state workers. Again, we see that local shocks can have large impacts for local workers while still generating an overall incidence that is spread widely.

Figure 15a (Table A20, col. 1-3) displays the share of all employment losses experienced by each employment status/initial earnings quartile. For disasters featuring 25% local job loss, nearly 35% of lost net employment is experience by those already unemployed, with the share falling monotonically from 27.8% to 8.9% as one moves from the 1st to the 4th initial earnings quartile. Thus, high skilled workers seem relatively well insulated from employment losses, instead taking jobs from those at lower skill levels, creating a cascade of sorts. However, as the disaster becomes more intense, the burden of employment loss becomes more equally shared, with only 31% accounted for by initially nonemployed, and 11% accounted for by those initially in the highest earnings quartile.

However, Figure 16a (Table A21, Col. 1-6), which examines employment incidence by distance and initial earnings jointly, paints a richer picture. Among those initially employed in the focal tract, the increase in the probability of unemployment from the least severe (25%) disaster is actually larger for employed workers than for initially unemployed workers: younger (older) initially unemployed workers experience a 0.2 (0.6) percentage point increase in destination unemployment, while workers at initial earnings quartiles 1-4 experience increases in unemployment rate of 4.4, 3.6, 2.8 and 2.2 percentage points, respectively. This is primarily due to the fact that initially unemployed workers had the least to lose: they were fairly likely to be unemployed again in the absence of a disaster. However, among workers a tract away or further, the employment losses are greatest among the existing unemployed. As the disaster becomes more severe, this pattern becomes even more pronounced. For the most severe (100% job loss) disaster, initially unemployed older (younger) local workers experience a 0.6 (1.6) percentage point increase in unemployment rate, while local 1st and 4th earnings quartile workers experience 21.8 and 19.4 percentage point increases, respectively (Figure 16b (Table A21, Col. 7-12)).
When averaging across distance categories, welfare losses (Figure 15b, Table A20, col. 4-6) are more concentrated among the initially employed and particularly higher paid, and are nearly unaffected by the severity of the disaster: unemployed workers account for only around 9% of welfare incidence, and the share of welfare losses falling on initially employed workers increases monotonically from 21% for the 1st earnings quartile to 25.5% for the 4th regardless of severity. As with employment incidence, however, these numbers obscure substantial variation in the relative skill incidence of disasters by distance from the focal site. For disasters involving a 25% job loss, workers 1 state away experience utility losses, relative to the least affected origin worker type, equivalent to $-118 in annual earnings regardless of initial skill (Figure 17a, Table A22). The values are between $-136 and $-139 for the 100% job loss disasters (table available upon request). Differences in welfare losses are similarly small for all distance bins except workers initially employed (or unemployed) in the focal tract. However, both younger and older workers initially unemployed within the focal tract are predicted to lose the equivalent of around $300 in utility in the 25% disaster, while workers in earnings quartiles 1-4 are predicted to lose $-4,755, $-6,043, $-7,022, and $-7,531, respectively. For the disasters featuring 100% local job loss, these values rise to around $-600 and $-13,950, $-18,000, $-21,550, and $-23,550 respectively. Thus, welfare losses are particularly large among local high skilled workers (who had the most to lose), although possibly smaller as a share of initial utility from annual earnings.

Finally, while quantifying the employment and utility incidence of disasters is important for allocating relief funds, policymakers and local communities are also worried about being inundated by flows of migrants away from disaster sites. Thus, Figure 18a (Table A23) displays, for each disaster intensity, the change in the probability of being employed at establishments in each distance bin relative to the focal tract for workers who were initially employed (or unemployed) in the census tract hit by the natural disaster. First, note that for the mildest disaster, the decrease in within-tract employment for workers initially at the focal tract is only 10.6%, despite a 25% overall decrease in local positions. This is in part because many of these workers would have moved to jobs away from the tract even in the absence of the shock, but also because existing local workers are able to retain a greater share of the jobs that remain. Even when all local jobs are lost, the decrease in the share of local workers staying in the tract is only 70.5%, revealing that 29.5% of such workers would have moved out in absence of the disaster. Adjacent tracts absorb an extra 0.5%, 1.2%, and 3.0% of workers initially in the focal tract, respectively, in the 25%, 50%, and 100% job loss scenarios, relative to a counterfactual in which no disaster occurs. Overall, an additional 1.9%, 4.5%, and 11.6% of the workers initially in the focal tract end up employed in other tracts within the original PUMA after the 25%, 50%, and 100% job loss scenarios. Locations outside the PUMA but within the state take on an additional 4.2%, 10.4%, and 27.6% of those initially employed in the focal tract in the three disaster scenarios, while locations outside the state take on an additional 0.6%, 1.6%, and 4.3% in the three scenarios, with
the remaining share of workers experiencing unemployment. Thus, while a relatively small share of employees find employment nearby, this share increases in the degree of initial displacement.

Figure 18c (Table A24) displays separate distributions of destination employment locations for target tract workers in each employment status/initial earnings quartile. Even in the most severe disaster, only an additional 0.9% (1.7%) of younger initially unemployed workers move away from the focal tract, relative to the counterfactual. The few that would have gotten local jobs remain unemployed instead. By contrast, the share moving to nearby locations is much larger for initially employed workers, and is somewhat consistent across initial earnings quartiles. Since most high paid workers are retained or continue to work nearby in the absence of the shock (only 12.4% would have transitioned away from the focal tract), the extreme 100% disaster engenders a particularly large mobility response for such workers: an additional 12.6% move to another tract within the PUMA (relative to the counterfactual), an additional 38.2% move to a different PUMA within the state, and an additional 7.1% move to a different state. For workers initially in the 1st earnings quartile, who were more likely to switch jobs in the absence of the disaster, the corresponding increases are only 11.2%, 24.3%, and 2.6%, respectively. Relative to lower skilled workers, the mobility response for initially high earning workers is disproportionately muted for lesser disasters, though, because they are better able to capturing the remaining local jobs than less skilled workers (Figure 18b, Table A24).

6.3 Model Validation

The simulations consider relatively large, locally focused labor demand shocks, but the estimated surplus parameters \( \hat{\Theta}^{D'in-D} \) that underlie them are identified from millions of quotidian worker job transitions stemming from natural job turnover and preference or skill changes over the life cycle that generate huge amounts of offsetting churn in the American labor market. Thus, one might reasonably wonder whether the parameters that govern ordinary worker flows remain valid when considering the response to sizable, locally targeted positive or negative shocks. To address this concern, a model validation exercise was performed in which parameters from a two-sided assignment model estimated on pre-shock ordinary worker flows were used to forecast the reallocation of workers that followed actual local economic shocks observed in the LEHD sample. I evaluate model fit using an index of dissimilarity between the predicted and actual transition group distributions \( P(g) \), and average this index across 514 shocks defined by census tract/year combinations featuring large positive or negative changes in employment (at least 100 workers and at least 10% of the pre-shock employment level in the tract), without offsetting contemporaneous shocks to other tracts in the same PUMA or shocks in other years within the same tract. Appendix A7 describes the exercise in detail, while Appendix Table A25 reports the results.
To summarize, on average only 7% of transitions nationally would need to be reallocated across transition groups in order for the two-sided assignment model to perfectly match the actual allocation that occurred following the shocks. Because transition groups feature locations that are much more narrowly defined within the PUMA associated with the shock (tracts rather than PUMAs or states), the model would need 36.5% of transitions originating in the targeted PUMA to be reallocated to other groups in order to perfectly match the true within-PUMA distribution $P(g)$. However, a substantial share of the “incorrect” predictions involve either slight differences in destination tract within the same distance bin from the targeted census tracts or slightly mismatched combinations of destination establishment characteristics (combos of size/avg. pay/supersector categories). When the group space is collapsed so that origin and destination locations are defined by 42 distance bins from the targeted tract and non-location establishment characteristics are excluded (though worker initial earnings categories remain), the share of transitions that must be reallocated across collapsed groups to perfectly fit the data falls to 0.9% nationally and 5.3% among transitions originating in the target PUMA. This is despite the fact that there remain 10,752 transition groups with only 294 restrictions imposed by the provided origin and destination type marginal PMFs. The model also fits well the origin type distribution of workers who end up or remain unemployed after the shock, particularly when the type space is collapsed to the distance bins I considered in the figures discussed above, suggesting that those counterfactual estimates of the share of employment incidence experienced by different skill/distance bin combinations are likely to be accurate.

Furthermore, the model vastly outperforms the predictions from an alternative one-sided parametric conditional logit model fit to the conditional choice probabilities (CCPs) $P(g|d)$ from the prior pair of years, suggesting that retaining the full information contained in the pre-year data rather than compressing to a smaller set of parameters has considerable value. The two-sided model also outperforms (though by a smaller margin) other alternative one-sided nonparametric forecasts that hold fixed the full set of either raw or smoothed CCPs, so that $P(g)^{y,CF} = h^y(d)P^{y-1}(g|d)$, suggesting that requiring market clearing does have additional predictive value.

Taken together, the model does quite a good job of predicting the reallocation of workers across job types and particularly across employment/unemployment status that follows major local labor market shocks.

7 Conclusion

Building on the approach of Choo and Siow (2006), this paper models the transition of the U.S. labor market across adjacent years as a large-scale assignment game with transferable utility, and uses a very large set of estimated parameters from the model to simulate the welfare incidence across locations and worker skill categories of a variety of alternative local
labor demand shocks designed to resemble different stimulus packages and natural disasters.

I show that one can still produce forecasts of welfare incidence on both sides of the market from changes in agent type composition on either side of the market even when singles are either not observed or observed on only one side of the market. By basing simulations on millions of composite joint surplus parameters rather than reducing the data to a much smaller set of fundamental utility or production function parameters, the “sufficient statistics” approach used here can fully exploit the massive scale of the administrative LEHD database to capture multidimensional heterogeneity on both sides of a two-sided market without placing unjustified structure on the job matching technology.

The method can be customized to forecast the incidence of any particular shock composition or magnitude in any location, and incidence can be determined across groups of agents on either side of the market defined by any arbitrary combination of observed characteristics, including categorical characteristics without a natural ordering such as race, industry or location. Given appropriate administrative matching data, the approach here could also be easily adapted to the student-college matching or patient-doctor matching contexts, among other applications.

I find that U.S. labor markets are still quite local, in that the per-worker welfare gains from a locally targeted labor demand shock are substantially larger for workers in the focal census tract than even workers one or two tracts away. Nonetheless, because the workers initially working within a very small radius of the local shock are such a small share of the entire U.S. labor force competing for positions, I also find that in most specifications greater than 40% of the welfare gain from a very local stimulus package, regardless of establishment composition, redounds to workers initially working out of state, with only about 7.0% of the welfare gains going to existing workers in the PUMA that contains the focal census tract.

I also document a high degree of heterogeneity in skill incidence among very local workers across demand shocks featuring different establishment size, average pay, and industry supersector composition, suggesting that the type of establishment targeted by a local development policy has major implications for the groups of workers most likely to benefit. That said, as these alternative shocks ripple across space through a chain of job transitions, their skill incidence becomes increasing similar, so that the overall skill composition of welfare gains across all workers (not just local workers) is extremely similar across different types of demand shocks. Thus, local policymakers need to be much more careful about which projects are subsidized than would state or national funders of local projects, since such funders would internalize the ripple effects.

Finally, I find that positive and negative shocks have asymmetric impacts on local employment, with negative shocks displaying a much greater geographic concentration of employment losses than the corresponding employment gains from positive shocks. This is because most workers would have been working anyway in the absence of a positive shock,
while most have jobs at risk when negative shocks occur.

Going forward, two extensions seem particularly worthwhile. First, following Caliendo et al. (2015), rather than computing incidence over a one year horizon, the assignment game could be made dynamic, and the time path of incidence of both temporary and permanent local shocks could be traced out. Second, following Galichon and Salanié (2015) and Chiappori et al. (2009), one could relax the assumption that the idiosyncratic job-match-level surplus shocks are i.i.d. across worker and position types so as to incorporate an even more flexible set of substitution patterns than those featured in the present model.

References


Chen, Liang (2017) ‘Compensation, moral hazard, and talent misallocation in the market for ceos’


Schmutz, Benoît, and Modibo Sidibe (2016) ‘Frictional labor mobility’

Shapley, LS, and Martin Shubik (1972) Game Theory in Economics: Chapter 3: the” rules of the Games” (Rand)


Figure 1: Empirical Distribution of 2010-2011 Job Transitions

Notes: The bar heights capture the shares of all worker transitions between dominant positions in 2010 and 2011 in which the geographic distance between these positions’ establishments fell into the distance bins indicated by the bar labels. “0/1/2/3+ Tct” indicates that the two establishments were in the same tract or one, two, or 3+ tracts away (by tract pathlength) within the same PUMA. “1/2/3+ PUMA” and “1/2+ State” indicate the PUMA pathlength (if within the same state) and state pathlength, respectively.
Figure 2: Distribution of the Distance Between Workers’ Origin Position and the Census Tract Targeted by the Simulated Stimulus Package: Average across All Simulated Stimuli

Notes: The bar heights capture the shares of all workers for whom the geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels (computed separately for each target tract, then averaged across all 500 target tracts). “0/1/2/3+ Tct” indicates that the origin establishment was in the target tract or was one, two, or 3+ tracts away (by tract pathlength) within the same PUMA. “1/2/3+ PUMA” and “1/2+ State” indicate the PUMA pathlength (if within the same state) and state pathlength, respectively.
Figure 3: Probability of Obtaining a Stimulus Job and Share of All Stimulus Jobs Obtained by Distance From Focal Tract: Average across All Simulated Stimuli

(a) Probability of Obtaining a Stimulus Job

(b) Share of All Stimulus Jobs

Notes: The bar heights in Figure 3a capture the average probability of obtaining a stimulus job among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels. These probabilities average across different initial earnings categories and across stimulus packages featuring different firm compositions. Figure 3b displays the share of all stimulus jobs that redounds to workers in the chosen distance bin. “0/1/2/3+ Tct” indicates that the origin establishment was in the target tract or was one, two, or 3 or more tracts away (by tract pathlength) within the same PUMA, “1/2/3+ PUMA” and “1/2+ State” indicate the PUMA pathlength (if within the same state) and state pathlength (if in different states), respectively.
Figure 4: Change in P(Employed) and Share of Additional Employment Obtained by Distance From Focal Tract: Average across All Simulated Stimuli

(a) Change in P(Employed)

(b) Share of All Additional Employment

Notes: The bar heights in Figure 4a capture the average change in employment probability among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels. These probabilities average across different initial earnings categories and across stimulus packages featuring different firm compositions. Figure 4b displays the share of additional employment generated by the stimulus that redounds to workers in the chosen distance bin. “0/1/2/3+ Tct” indicates that the origin establishment was in the target tract or was one, two, or 3 or more tracts away (by tract pathlength) within the same PUMA. “1/2/3+ PUMA” and “1/2+ State” indicate the PUMA pathlength (if within the same state) and state pathlength (if in different states), respectively.
Figure 5: Expected Welfare Changes and Share of Total Welfare Gains by Distance From Focal Tract: Average across All Simulated Stimuli

(a) Expected Welfare Changes ($)

(b) Share of Total Welfare Gains

Notes: The bar heights in Figure 5a capture the average welfare gain (scaled in $ of annual earnings) among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels. Averages are taken across different initial earnings categories and across stimulus packages featuring different firm compositions. Figure 5b displays the share of all welfare gains generated by the stimulus that redounds to workers in the chosen distance bin. “0/1/2/3+ Tct” indicates that the origin establishment was in the target tract or was one, two, or 3 or more tracts away (by tract pathlength) within the same PUMA. “1/2/3+ PUMA” and “1/2+ State” indicate the PUMA pathlength (if within the same state) and state pathlength (if in different states), respectively.
Figure 6: Change in P(Employed) and Expected Welfare Gains and Shares of Additional Employment by Distance From Focal Tract and Industry of Stimulus

(a) Change in P(Employed)

(b) Expected Welfare Gain ($)

Notes: The bar heights within a particular group in Figure 6a and 6b capture the average change in employment probability and average welfare gain (scaled in $ of annual earnings), respectively, among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins defined in Figure 3. Each group of bars displays average outcomes across distance bins for a set of simulations featuring stimulus packages consisting of positions at firms in the industry supersector given by the group label. Averages are taken across different initial earnings categories and across stimulus packages featuring different firm size/average pay compositions, as well as across 500 simulations featuring different targeted census tracts for each supersector/firm size/firm avg. pay combo. “R/W Trade” = Trade, Transportation and Utilities. “Other Services” includes repair and maintenance firms, personal and laundry services, and religious/civic/professional organizations and private households.
Figure 7: Change in P(Employed) and Shares of Additional Employment by Distance From Focal Tract and Firm Size/Firm Average Pay Composition of the Stimulus Package

(a) Change in P(Employed)

(b) Share of Additional Employment

Notes: The bar heights within a particular group in Figure 7a and 7b capture the average change in employment probability and share of all additional employment, respectively, among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins defined in Figure 3. Each group of bars displays average outcomes across distance bins for a set of simulations featuring stimulus packages consisting of positions at firms in the combination of firm size/firm avg. pay categories given by the group label. Averages are taken across different initial earnings categories and across stimulus packages featuring different industry super-sector compositions, as well as across 500 simulations featuring different targeted census tracts for each supersector/firm size/firm avg. pay combo. “Sm-Lo”: Establishments in the 1st quartile of establishment size (based on employment) and 2nd quartile of average pay. “Sm-Hi”: 1st size quartile, 4th pay quartile “Lg-Lo”: 4th size quartile, 2nd pay quartile. “Lg-Hi”: 4th size quartile, 4th pay quartile.
Figure 8: Shares of Additional Employment and of Total Utility Gains among Workers Initially Employed (or Nonemployed) at Different Initial Earnings Quintiles (or Nonemployment): Average across All Simulated Stimuli

(a) Share of Additional Employment

(b) Share of Total Utility Gains

Notes: The bar heights in Figure 8a and 8b capture the average share of all employment gains and of all welfare gains, respectively, among workers whose origin employment status fell into the employment status/earnings quartiles indicated by the bar labels. Averages are taken across different bins capturing the distance between the workers initial (or most recent) employment and the targeted tract receiving the stimulus, across stimulus packages featuring different industry supersector/firm size/firm avg. pay compositions, as well as across 500 simulations featuring different targeted census tracts for each stimulus composition. “Q1/Q2/Q3/Q4”: Workers whose pay at their dominant job in the origin year placed them in the 1st/2nd/3rd/4th Quartile of the national earnings distribution. “<25 NE/25+ NE”: Workers who were not employed in the origin year who were younger than 25 years of age/at least 25 years of age.
Figure 9: Change in P(Employed) and in Expected Utility Among Workers Initially Employed at Different Combinations of Initial Earnings/Employment Status and Distance from Focal Tract: Average across All Simulated Stimuli

(a) Change in P(Employed)

(b) Change in Expected Utility

Notes: The bar heights within a particular group in Figures 9a and 8b capture the average change in employment probability and share of all additional employment, respectively, among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins defined in Figure 3. Each group of bars displays average outcomes across distance bins for groups of workers defined by their origin employment status/earnings category. Averages are taken across stimulus packages featuring different firm supersector/size/avg. pay compositions, as well as across 500 simulations featuring different targeted census tracts for each firm composition. “Q1/Q2/Q3/Q4”: Workers whose pay at their dominant job in the origin year placed them in the 1st/2nd/3rd/4th Quartile of the national earnings distribution. “<25 NE/25+ NE”: Workers who were not employed in the origin year who were younger than 25 years of age/at least 25 years of age.
Figure 10: Expected Welfare Gains (in $) Among Workers Originally Working in the Targeted Tract by Initial Earnings/Employment Status: By Industry Supersector or Firm Size Quartile/Firm Pay Quartile Combination

(a) Expected Welfare Gain by Industry Supersector

(b) Expected Welfare Gain by Firm Size Quartile/Firm Avg. Pay Quartile Combination

Notes: The bar heights within a particular group in Figures 10a and 10b capture the average welfare gain from a 500 person stimulus package among workers whose origin employment status/earnings quartile fell into the bins defined in Figure 8. Each group of bars displays average outcomes among simulated stimulus packages featuring positions within the particular industry supersector (in Figure 10a) or particular firm size/firm avg. earnings quartile combination (in Figure 10b) given by the group’s label. Averages are taken across different initial distance from the focal tract of the shock, across stimulus packages featuring different industry supersector compositions (in Figure 10a) or different firm size/firm pay compositions (in Figure 10b), as well as across 500 simulations featuring different targeted census tracts for each supersector/firm size/firm avg. pay combo. See Figures 6 and 7 for expanded definitions of group labels.
Figure 11: Expected Utility Changes Among Workers Originally Working in the Targeted Tract by Initial Earnings/Employment Status: All Stimulus Packages

Notes: Each line traces the expected welfare gain generated by a stimulus package featuring 500 positions among firms with a particular supersector/firm size quartile/firm pay quartile combination across alternative origin nonemployment or earnings quartile categories. 32 different lines corresponding to 32 different firm supersector/size/pay level compositions are displayed. Averages are taken across different initial distance from the focal tract of the shock, as well as across 500 simulations featuring different targeted census tracts for each supersector/firm size/firm avg. pay combo. “Q1/Q2/Q3/Q4”: Workers whose pay at their dominant job in the origin year placed them in the 1st/2nd/3rd/4th Quartile of the national earnings distribution. “<25 NE/25+ NE”: Workers who were not employed in the origin year who were younger than 25 years of age/at least 25 years of age.
Figure 12: Heterogeneity in Several Incidence Measures by Distance from Focal Tract Across Focal Tracts of Varying Population and Employment Size

(a) Change in \(P(\text{Employed})\)  

(b) Share of Employment Gains  

(c) Avg. Welfare Gain (\$)  

(d) Share of Welfare Gains

Notes: The bar heights within a particular group in Figures 12a-12d capture the average measure of stimulus incidence associated with the chosen figure from a 500 person stimulus package among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins defined in Figure 3. Each group of bars displays this incidence distribution across distance bins for particular subset (indicated by the group’s label) of the 500 simulations featuring different focal tracts that were performed. In addition to averaging over the simulations featuring different target tracts within the chosen subset, the displayed results also average over different stimuli featuring the same target census tract but different firm compositions. “All”: Average is taken among all 500 target tracts. “Rural”/“Urban”: Average is taken among the 100 target tracts with the smallest/largest residential population density. “Lo Emp”/“Hi Emp”: Average is taken among the 100 target tracts with the smallest/largest numbers of destination jobs located within the tract (in the absence of a counterfactual shock).
Figure 13: Assessing the Value of Restricting Stimulus Jobs to Fill Positions Within the Target PUMA: Spatial Employment and Welfare Incidence for Restricted and Unrestricted Stimulus Packages (Each Featuring 500 Positions at a Large Low-Paying Manufacturing Firm)

(a) Change in P(Employed)
(b) Share of Employment Gains
(c) Avg. Welfare Gain ($)
(d) Share of Welfare Gains

Notes: The bar heights capture the average measure of stimulus incidence associated with the chosen figure from a 500 person stimulus package among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the labels. The thin, light blue bars capture the case in which the new positions are restricted to be filled by existing workers within the targeted PUMA, while the wide, dark blue bars. Each bar represents an average over 500 simulations featuring different target census tracts as well as over 32 packages for each these 500 simulations featuring different firm composition (combinations of industry supersector, firm size quartile, and firm avg. pay quartile). “0/1/2/3+ Tct” indicates that the origin establishment was in the target tract or was one, two, or 3 or more tracts away (by tract pathlength) within the same PUMA. “1/2/3+ PUMA” and “1/2+ State” indicate the PUMA pathlength (if within the same state) and state pathlength (if in different states), respectively.
Figure 14: Employment and Welfare Incidence from a Natural Disaster by Distance From Focal Tract and Severity of the Disaster (25%/50%/100% Jobs Lost)

(a) Change in $P(\text{Unemployed})$  
(b) Share of Employment Losses  
(c) Expected Welfare Loss  
(d) Share of Total Welfare Losses

Notes: The bar heights within a particular group in Figures 14a-14d capture the average value of the incidence measure associated with the figure from a set of simulated natural disasters among workers whose geographic distance between their origin establishments and the census tract experiencing the disaster fell into the distance bins defined in Figure 3. Each group of bars displays the distribution of disaster losses across distance bins for a set of simulations featuring different shock intensities: either 25%, 50% or 100% of the original positions in the focal census tract are eliminated. For each disaster intensity, averages are taken across 500 simulations featuring different targeted census tracts.
Figure 15: Expected Shares of Additional Nonemployment and Welfare Losses Produced by a Natural Disaster among Workers Initially Employed (or Nonemployed) at Different Initial Earnings Quintiles (or Nonemployment), by Disaster Severity (25%/50%/100% of Jobs Lost)

Notes: The bar heights within a particular group in Figures 15a and 15b capture the average share of additional nonemployment and welfare losses, respectively, from a set of simulated natural disasters among workers whose origin employment status/earnings quartile fell into the bins defined in Figure 8. Each group of bars displays the distribution of disaster losses across employment status/earnings quartile bins for a set of simulations featuring different shock intensities: either 25%, 50% or 100% of the original positions in the focal census tract are eliminated. For each disaster intensity, averages are taken across 500 simulations featuring different targeted census tracts.
Figure 16: Change in P(Unemployed) and Share of Total Employment Losses Produced by a Natural Disaster (25% or 100% of Jobs Lost) for Randomly Chosen Workers Initially Employed at Different Combinations of Initial Earnings/Employment Status and Distance from Focal Tract

(a) Change in P(Unemployed), 25% of Jobs Lost
(b) Change in P(Unemployed), 100% of Jobs Lost
(c) Share of Total Employment Losses, 25% of Jobs Lost
(d) Share of Total Employment Losses, 100% of Jobs Lost

Notes: The bar heights within a particular group in Figures 16a-16d capture either the average change in unemployment probability or the share of all employment losses (depending on the figure) from a set of simulated natural disasters among workers whose geographic distance between their origin establishments and the census tract experiencing the disaster fell into the distance bins defined in Figure 3. In Figures 16a and 16c, 25% of jobs in the targeted census tract are eliminated, while in Figures 16b and 16d 100% are eliminated. Each group of bars displays the distribution of losses across distance bins for groups of workers defined by their origin employment status/earnings category. For each disaster intensity, averages are taken across 500 simulations featuring different targeted census tracts. “Q1/Q2/Q3/Q4”: Workers whose pay at their dominant job in the origin year placed them in the 1st/2nd/3rd/4th Quartile of the national earnings distribution. “<25 NE/25+ NE”: Workers who were not employed in the origin year who were younger than 25 years of age/at least 25 years of age.
Figure 17: Expected Welfare Loss (in $) and Share of Total Welfare Losses Produced by a Natural Disaster (25% or 100% of Local Jobs Lost) for Randomly Chosen Workers Initially Employed at Different Combinations of Initial Earnings/Employment Status and Distance from Focal Tract

(a) Expected Welfare Loss, 25% of Jobs Lost
(b) Expected Welfare Loss, 100% of Jobs Lost
(c) Share of Total Welfare Losses, 25% of Jobs Lost
(d) Share of Total Welfare Losses, 100% of Jobs Lost

Notes: The bar heights within a particular group in Figures 16a-16d capture either the average welfare loss (scaled in $ of annual earnings) or the share of all welfare losses (depending on the figure) from a set of simulated natural disasters among workers whose geographic distance between their origin establishments and the census tract experiencing the disaster fell into the distance bins defined in Figure 3. In Figures 17a and 17c, 25% of jobs in the targeted census tract are eliminated, while in Figures 17b and 17d 100% are eliminated. Each group of bars displays the distribution of losses across distance bins for groups of workers defined by their origin employment status/earnings category. For each disaster intensity, averages are taken across 500 simulations featuring different targeted census tracts. “Q1/Q2/Q3/Q4”: Workers whose pay at their dominant job in the origin year placed them in the 1st/2nd/3rd/4th Quartile of the national earnings distribution. “<25 NE/25+ NE”: Workers who were not employed in the origin year who were younger than 25 years of age/at least 25 years of age.
Figure 18: Changes in the Distribution of Employment Locations (or Nonemployment) for Workers Initially Employed in the Focal Tract after a Natural Disaster (25%, 50%, or 100% of Jobs Lost)

(a) 25/50/100% of Jobs Lost, at Different Distances from the Focal Tract (Averaging Across Initial Earnings/Employment Statuses)

(b) 25% of Jobs Lost, at Different Distances from the Focal Tract, by Initial Earnings/Employment Status

(c) 100% of Jobs Lost, at Different Distances from the Focal Tract, by Initial Earnings/Employment Status

Notes: The bar heights within a particular group in Figures 18a-18c capture the impact of experiencing a natural disaster that removes either 25%, 50%, or 100% of jobs in the focal tract on the probability that a worker initially employed (or most recently employed) in the targeted tract would be employed at a position whose geographic distance from the census tract experiencing the disaster fell into the distance bins defined in Figure 3 (or become/remain unemployed). In Figure 18a, each group of bars displays the change in destination employment probabilities for a particular disaster intensity (25%, 50% or 100% of jobs lost in the target tract), and plotted values are averages over different initial employment status/earnings quartile categories. In Figures 18b and 18c, each group of bars displays the change in destination employment probabilities for groups of workers defined by their origin employment status/earnings category. For each disaster intensity, averages are taken across 500 simulations featuring different targeted census tracts. “Q1/Q2/Q3/Q4”: Workers whose pay at their dominant job in the origin year placed them in the 1st/2nd/3rd/4th Quartile of the national earnings distribution. “<25 NE/25+ NE”: Workers who were not employed in the origin year who were younger than 25 years of age/at least 25 years of age.
Appendix

A1 Proof of Proposition A1

Proposition A1:

Suppose the following assumptions hold:

1’) The assumptions laid out in sections 2.3 and 3 continue to hold. Namely, each joint surplus $\pi_{ijk}$ is additively separable in the group-level and idiosyncratic components, the vector of idiosyncratic components $\epsilon_{ij(i)k}$ is independently and identically distributed, and follows the type 1 extreme value distribution, and Assumptions 1 and 2 hold.

2’) The set of destination positions $k \in \tilde{K}$ that will be filled in the stable counterfactual assignment are known in advance, and the set of destination positions $k \in \tilde{K}$ that will remain unfilled in the stable counterfactual assignment are ignorable, in the sense that their existence does not change the assignment nor the division of surplus among the remaining set of positions $K$ and set of workers $I$.

3’) $\frac{1}{|g|} \sum_{k:g(i,j(i),k)=g} e^{-\frac{q_k}{\sigma}} \approx \frac{1}{|d|} \sum_{k:d(k)=d(g)} e^{-\frac{q_k}{\sigma}} = C_d(g) \forall (g,i)$.

4’) $P(g|i,d(g)) \approx P(g|o(g),d(g)) \forall (g,i)$.

Then the group-level assignment $P^{CF}(g)$ that satisfies the following $O-1$ excess demand equations represents the unique group-level equilibrium assignment $P^{CF*}(g)$ consistent with the unique worker/position level stable matching $\mu^{CF}$:

$$\sum_{d \in D} h^{CF}(d) \sum_{g:o(g)=2} P^{CF}(g|d,C^{CF}_2,\ldots,C^{CF}_O) = f^{CF}(2)$$

$$\vdots$$

$$\sum_{d \in D} h^{CF}(d) \sum_{g:o(g)=O} P^{CF}(g|d,C^{CF}_2,\ldots,C^{CF}_O) = f^{CF}(O)$$

(25)

where $P^{CF}(g|d,C^{CF}_2,\ldots,C^{CF}_O)$ is given by:

$$P^{CF}(g|d) = \frac{e^{\frac{\theta^{CF}}{\sigma} \sum_{o(g)} f^{CF}(o(g)) C^{CF}_o}}{\sum_{o' \in O} \sum_{g' \in (o,d)} e^{\frac{\theta^{CF}}{\sigma} \sum_{o'(g')} f^{CF}(o(g')) C^{CF}_{o'}}} \forall d \in [1,\ldots,D]$$

(26)
Proof: Proposition A1 states that assignment $P^{CF}(g)$ implied by the vector of mean utility values $C^{CF} = [1, C_2, \ldots, C^{CF}_O]$ that solves the system of equations (25) in fact represents the unique group-level stable (and equilibrium) assignment $P^{CF*}(g)$.

First, note that if unfilled positions are ignorable for the counterfactual assignment, then we can focus on finding a stable assignment of a restricted version of the assignment game in which only remaining $K$ positions need to be considered. As discussed in footnote 19, Assumption 2' implicitly requires that no position that remains unfilled is ever the second-best option for any worker who takes a job in the destination period.

Furthermore, Assumption 2' imposes that each of the remaining positions will be filled in any stable matching. Recall that stability in the individual-level matching $\mu^{CF}$ requires:

$$\mu^{CF}_{ij(k)} = 1 \iff k \in \arg \max_{k \in K \cup 0} \pi_{ij(k)} - q^C_k$$ and $i \in \arg \max_{i \in I \cup 0} \pi_{ij(k)} - r^C_i$ \hspace{1cm} (27)

Assumption 2' allows us to replace $i \in \arg \max_{i \in I \cup 0} \pi_{ij(k)} - r^C_i$ with $i \in \arg \max_{i \in I} \pi_{ij(k)} - r^C_i$. In other words, I assume in advance that the individual rationality conditions that any proposed match yield a higher payoff to the position than remaining vacant, $\pi_{ij(k)} - r_i > \pi_{0k}$ when $\mu_{ik} = 1$, are satisfied and can be ignored. Implicitly, this requires that the joint surpluses to workers and firms from matching up are sufficiently large relative to both workers’ and firms’ outside options.\footnote{This implicitly requires that the unobserved draws $\epsilon_{0k}$ for position vacancy values are taken from a bounded distribution rather than the Type 1 extreme value distribution.}

Imposing Assumption 2' will probably cause utility losses among local workers from negative local labor demand shocks to be overstated, since some workers would likely find jobs at positions that were not willing to hire at the original wage level but would enter the labor market at lower wage levels. Conversely, gains to local workers from positive shocks may be understated, since some local firms that filled positions at the original wage levels might choose to remain vacant (or move to other locations) when competition for local workers becomes more fierce.

In our applications the number of positions that will be filled (which Assumption 2' imposes will be known) is greater than the number of workers seeking positions ($I$). In order to be able to consistently allocate workers to transition groups, even when they move to (or remain in) nonemployment, I define a “nonemployment” destination type as the last destination type $D$. Because the number of workers who end up unemployed is assumed to be known, I allocate enough “nonemployment” positions within type $D$, $h^{CF}(D)$, so that the number of workers $I$ equals the number of “positions” $K$, once $K$ includes the dummy nonemployment positions. I then normalize this common number of workers and firm positions (assumed to be very large) to be 1, and reinterpret $f^{CF}(o)$ and $h^{CF}(d)$ as probability mass functions providing shares of the relevant worker and position populations rather than counts.

As discussed in Section 3, Assumption 1', when combined with the stability conditions
implies that the probability that a given position $k$ will be filled by a particular worker $i$ is given by the logit form (9). When combined with Assumptions 1 and 2 (also cited by Assumption 1'), this implies that the group-level conditional choice probability $P(g|d)$ takes the form (26) for any destination types $d$ that are composed of positions $k$ (as derived in section 3).

However, note the statement of Proposition A1 makes it clear that the form (26) also holds for the last type $D$, which contains the “dummy” nonemployment positions whose “choices” will be workers moving to nonemployment. The stability conditions (27) do not provide any justification for why these dummy nonemployment positions should be filled via the same logit form as the other destination types that consist of actual positions at firms. Thus, the inclusion of these dummy positions, and the assumption that the probability distribution over alternative groups representing different worker and job match characteristics $(o(g), z(g))$ follows the logit form, are mere computational devices to calculate the equilibrium assignment. That this computational device in fact yields the unique stable assignment for the proposed counterfactual labor market is the primary reason Proposition A1 requires a proof.

However, the stability conditions and Assumption 1’ imply that the probability that a given worker $i$ will choose a particular position $k$ (where $k = 0$ represents nonemployment) is also given by the logit form (Decker et al. (2013)):

$$P_{CF}(k|i) = \frac{e^{\theta_{CF}g - q_{CF}k}}{\sum_{k' \in K \cup 0} e^{\theta_{CF}g - q_{CF}k'}}$$

(28)

This can then be aggregated (using the same steps as in section 3) to provide an expression for the probability that a randomly chosen worker from a given origin type $o$ matches with a position that yields a transition assigned to group $g$:

$$P_{CF}(g|o) = \frac{1}{|o|} \sum_{i \in o} \left( e^{\theta_{CF}g - q_{CF}k} \right) \left( \sum_{k' \in \{(i,j(i),k) = g e^{\theta_{CF}g - q_{CF}k'}} \right)$$

$$\sum_{k' \in K \cup 0} e^{\theta_{CF}g - q_{CF}k'}$$

(29)

Assumptions 3’ and 4’, which are analogues to Assumptions 1 and 2 in section 3, allow us to simply this expression to the following:

$$P_{CF}(g|o) = \frac{e^{\theta_{CF}g} \tilde{S}_{g|o(d),d} \tilde{h}_{CF}(d(g)) \tilde{C}_{CF}}{\sum_{d' \in D} \sum_{g' \in \{(o,d') e^{\theta_{CF}g'} g|o(d'),d} \tilde{h}_{CF}(d'(g')) \tilde{C}_{CF}}} \forall o \in [1, \ldots, O]$$

(30)

Assumption 3’ states that the discounted profits of alternative positions $k$ assigned to the same destination type $d$ are roughly the same. This implies that the profit share that
workers must provide to the position in a stable matching is approximately the same for their existing positions as for other positions in the same local area featuring the same industry and establishment size and establishment average pay categories, and can be summarized by a parameter $C_{d}^{CF}$ that is defined at the destination-type level.

Taken literally (given the characteristics I use to define groups), Assumption 4' states that every worker assigned to the same origin type starts the year in firms with the same number of destination positions, which clearly does not hold. More broadly, though, Assumptions 3' and 4' allow us to replace the term $\sum_{k,j}^{g_{i,j,k}=g} e^{-\frac{CF}{\sigma}}$ that depends on the individual $i$ with an expression $P^{CF}(g|o,d(g))h^{CF}(d(g))\tilde{C}_{d}^{CF}$ that depends on only group and destination-type level terms. Essentially, I am assuming that ignoring within-origin type variation in the number of positions at which they would be stayers (due to different establishment sizes of initial job matches) when aggregating is not generating significant bias in the counterfactual assignment and incidence estimates.\footnote{Note also that assumptions 3' and 4' may not be necessary conditions for Proposition A1. I am seeking an alternative proof of Proposition A1 that does not require assumptions 3' and 4'.}

Under Assumptions 1’ through 4’, the group-level stable matching must satisfy the following market clearing conditions, which specify that supply must equal demand for each destination position type $d$:

$$\sum_{o \in O} f^{CF}(o)(\sum_{g,d(g)=2} P^{CF^{*}}(g|o,\tilde{C}^{CF}) = h^{CF}(2) \quad (31)$$

$$\vdots$$

$$\sum_{o \in O} f^{CF}(o)(\sum_{g,d(g)=D} P^{CF^{*}}(g|o,\tilde{C}^{CF}) = h^{CF}(D) \quad (33)$$

where $\tilde{C}^{CF}$ represents the $D - 1$ length vector $[1, \tilde{C}^{CF}_{2}, \ldots, \tilde{C}^{CF}_{D}]$ and each conditional probability $P^{CF^{*}}(g|o,\tilde{C}^{CF})$ takes the form in (30).

Note in particular that Assumption 2’ allows us to ignore the possibility that supply might exceed demand for some destination position types (implying some vacant positions). In this alternative position-side system of equations, the expressions for each conditional probability $P^{CF^{*}}(g|o)$ do in fact stem directly from the necessary stability conditions. And all of the feasibility conditions for a stable matching are incorporated into the zero-excess demand equations (since $P^{CF^{*}}(g|o)$ sum to 1 by construction, the assignment $P^{CF^{*}}(g)$ that satisfies this system necessarily sums to the origin-type PMF $f^{CF}(o)$). Thus, the proof by Decker et al. (2013) that there exists a unique group-level assignment that satisfies all of the group-level feasibility and stability conditions (and is thus consistent with a stable matching in the assignment game defined at the level of worker-position matches) applies here.

If one wished, one could directly compute the unique group-level counterfactual assign-
ment \( P^{CF^*}(g|o) \) by finding a \( D - 1 \) length vector \( \tilde{C}^{CF} \) that solved this system, and constructing the implied assignment by plugging this vector into the conditional probability expressions (30). However, when \( D \gg O \), solving this system is considerably more computationally burdensome than solving the worker-side counterpart (25), which features \( O - 1 \) equations. Thus, the remainder of this proof is devoted to showing that any assignment \( P^{CF}(g) \) implied by a solution to (25) must equal the assignment \( P^{CF^*}(g) \) implied by a solution to (33). And since we know that the latter solution represents the unique group-level matching consistent with stability in the assignment game, the former solution must also be unique, and must also represent the group-level matching consistent with stability in the assignment game. Essentially, this amounts to showing that the device of adding “dummy” nonemployment positions present in (25) appropriately incorporates the surpluses \( \pi_{i0} \) that workers obtain from staying single.

Consider an \( O \) length vector \( C^{CF} = [1, C_2^{CF}, \ldots, C_O^{CF}] \) that solves (25) and generates assignment \( P^{CF}(g) \). I wish to show that one can use \( C^{CF} \) to construct an alternative \( D \) length vector \( \tilde{C}^{CF} = [1, \tilde{C}_2^{CF}, \ldots, \tilde{C}_D^{CF}] \) that solves (33), and that the assignment it generates, \( P^{CF^*}(g) \), equals \( P^{CF}(g) \).

I propose the following vector \( \tilde{C}^{CF} \):

\[
\tilde{C}_d^{CF} = \frac{\sum_{o=1}^{O} \sum_{g'(o,g'),d(g')} \frac{\theta_{g'}}{\sigma} f^{CF}(o) \mathbb{S}_{g'|o,D} \tilde{C}_o^{CF} \delta_{g'|o,D}}{\sum_{o=1}^{O} \sum_{g'(o,g'),d(g')} \frac{\theta_{g'}}{\sigma} f^{CF}(o) \mathbb{S}_{g'|o,d} C_o^{CF}} \quad \forall \ d \in [1, \ldots, D] \quad (34)
\]

Here, the numerator captures the inclusive value (as defined by Menzel (2015)) associated with the nonemployment destination type \( D \), while the denominator captures the inclusive value for the chosen destination type \( d \). This implies that \( \tilde{C}_D^{CF} = 1 \). While any destination type could be chosen as the one whose mean exponentiated profit value is normalized, normalizing the nonemployment type is particularly appealing, since it implies “profit” values of 0 for the dummy nonemployment destination type \( D \) (\( \tilde{C}_D^{CF} = e^{\tilde{g}_D} = e^0 = 1 \)).

To conserve notation, let \( \lambda \) represent the inclusive value associated with the nonemployment destination type \( D \), the numerator in (39):

\[
\lambda = \sum_{o=1}^{O} \sum_{g'(o,g'),d(g')} \frac{\theta_{g'}}{\sigma} f^{CF}(o) \mathbb{S}_{g'|o,D} \tilde{C}_o^{CF} \quad (35)
\]

Note that \( \lambda \) is independent of destination type.

I begin by showing that the assignments implied by the vectors \( [C_1^{CF}, \ldots, C_O^{CF}] \) and \( [\tilde{C}_1^{CF}, \ldots, \tilde{C}_D^{CF}] \) are identical: \( P^{CF}(g) = P^{CF^*}(g) \).

Note first that since \( C^{CF} \) solves the worker-side system of excess demand equations (25),
I know that

\[
\sum_{d' \in D} h_{d'}^{CF} \sum_{g' \in \{o, d', o', \ldots\}} e^{\theta h_{d'}^{CF}} \sum_{g'=1}^{O} \sum_{g':(o(g'),d(g'))=(o',d')} e^{\frac{\theta}{\sigma}} f_{g'}^{CF}(o') \sigma_{g' | o,d'} C_{g'}^{CF} = f_{d'}^{CF}(o) \forall o \in [1, O]
\]

\[
\Rightarrow \sum_{d' \in D} \sum_{g' \in \{o, d', o', \ldots\}} e^{\frac{\theta}{\sigma}} \sum_{g'=1}^{O} \sum_{g':(o(g'),d(g'))=(o',d')} e^{\frac{\theta}{\sigma}} f_{g'}^{CF}(o') \sigma_{g' | o,d'} C_{g'}^{CF} = \frac{1}{C_{o}^{CF}} \forall o \in [1, O]
\]

\[
\Rightarrow \sum_{d' \in D} \sum_{g' \in \{o, d', o', \ldots\}} e^{\frac{\theta}{\sigma}} \sum_{g'=1}^{O} \sum_{g':(o(g'),d(g'))=(o',d')} e^{\frac{\theta}{\sigma}} f_{g'}^{CF}(o') \sigma_{g' | o,d'} C_{g'}^{CF} = \frac{\lambda}{C_{o}^{CF}} \forall o \in [1, O]
\]

(36)

I can now proceed:

\[
P_{d'}^{CF} (g) = f_{d'}^{CF}(o) P_{d'}^{CF} (g|o) = f_{d'}^{CF}(o) e^{\theta h_{d'}^{CF} C_{o}^{CF}} \sum_{d' \in D} \sum_{g' \in \{o, d', o', \ldots\}} e^{\frac{\theta}{\sigma}} \sum_{g'=1}^{O} \sum_{g':(o(g'),d(g'))=(o',d')} e^{\frac{\theta}{\sigma}} f_{g'}^{CF}(o') \sigma_{g' | o,d'} C_{g'}^{CF}
\]

\[
= f_{d'}^{CF}(o) e^{\theta h_{d'}^{CF} C_{o}^{CF}} \sum_{d' \in D} \sum_{g' \in \{o, d', o', \ldots\}} e^{\frac{\theta}{\sigma}} \sum_{g'=1}^{O} \sum_{g':(o(g'),d(g'))=(o',d')} e^{\frac{\theta}{\sigma}} f_{g'}^{CF}(o') \sigma_{g' | o,d'} C_{g'}^{CF}
\]

\[
= h_{d'}^{CF}(d) \sum_{d' \in D} \sum_{g' \in \{o, d', o', \ldots\}} e^{\frac{\theta}{\sigma}} f_{g'}^{CF}(o') \sigma_{g' | o,d'} C_{g'}^{CF}
\]

(37)

It remains to show that the chosen \(\bar{C}_{d'}^{CF}\) vector (39) solves (33). Consider the left-hand side of the excess demand equation for an arbitrarily chosen destination type \(d\) in the system.
(33). One can write:

\[
\begin{align*}
\sum_{o=1}^{O} \sum_{g:o(g),d(g)=(o,d)} f^{CF}(o) \cdot P^{CF^*}(g|o, \Theta^{CF}, \tilde{C}^{CF}) \\
\sum_{o=1}^{O} \sum_{g:o(g),d(g)=(o,d)} h^{CF}(d) \cdot P^{CF}(g|d, \Theta^{CF}, C^{CF}) \\
&= h^{CF}(d) \sum_{g:d(g)=d} P^{CF}(g|d, \Theta^{CF}, C^{CF}) \\
&= h^{CF}(d) \sum_{o=1}^{O} \sum_{g:o(g),d(g)=(o,d)} P^{CF}(g|d, \Theta^{CF}, C^{CF}) \\
&= h^{CF}(d)
\end{align*}
\]

where the last line uses the fact that \( P^{CF}(g|d) \) is a (conditional) probability distribution and thus sums to one. Since I have proved that the implied “demand” by workers for positions of an arbitrary destination type equals the “supply” \( h^{CF}(d) \), I have thus proved that \( \tilde{C}^{CF} \) solves the system (33).

Notice that the expression for the proposed equilibrium mean ex post profit vector (39) has value beyond its use in proving proposition A1. Once the \( O \)-vector of mean ex post utilities \( \{C^{CF}_o\} \) for each origin type have been computed, one can use (39) to directly calculate the mean ex post profit vector for each destination position type \( d \) without having to solve a system of \( D - 1 \) equations. This is quite valuable when \( D \gg O \), as it is in our application. Of course, the equivalent mapping can be inferred by symmetry for the opposite case where \( O \gg D \):

\[
C^{CF}_o = \frac{\sum_{d=1}^{D} \sum_{g^d:o(g^d),d(g^d)=(o,d)} \rho_{g^d} h^{CF}(d) S_{g^d|O,d} \tilde{C}^{CF}_d}{\sum_{d=1}^{D} \sum_{g^d:o(g^d),d(g^d)=(o,d)} \rho_{g^d} h^{CF}(d) S_{g^d|O,o} \tilde{C}^{CF}_d} \quad \forall o \in [1, \ldots, O]
\]

In section 3.2 I showed that these vectors are sufficient to determine both the worker and position type-level incidence of any counterfactual shocks to the composition or spatial distribution of labor supply and/or labor demand. Thus, at least in cases where the proposed model is a reasonable approximation of the functioning of the labor market (and housing supply is sufficiently elastic and agglomeration effects and other product market spillovers are second order), a proper welfare analysis of such shocks only requires solving at most \( \min\{O, D\} \) non-linear excess demand equations. Since an analytical Jacobian can be derived and fed as an input to non-linear equations solvers, relatively large scale assignment problems featuring thousands of types on one side of the market (and perhaps more on the opposite side) can be solved within a matter of minutes.
A2 Proof of Proposition 1

Proposition 1:

Define the set \( \Theta_{D^{-in-D}} \equiv \{ (\theta_{g'g}, \theta_{g''g''}) : (g, g', g'', g''') : o(g) = o(g''), o(g') = o(g'''), d(g') = d(g''), d(g'') = d(g''') \} \). Given knowledge of \( \Theta_{D^{-in-D}} \), a set \( \tilde{\Theta} = \{ \tilde{\theta}_g \} \) can be constructed such that the unique group level assignment \( P_{CF}(g) \) that satisfies the system of excess demand equations (24) using \( \theta_{g}^{CF} = \tilde{\theta}_g \forall g \) and arbitrary marginal distributions for origin and destination types \( f_{CF}(\*) \) and \( g_{CF}(\*) \) will also satisfy the corresponding system of excess demand equations using \( \theta_{g}^{CF} = \tilde{\theta}_g \forall g \) and arbitrary distributions \( f_{CF}(\*) \) and \( g_{CF}(\*) \). Furthermore, denote by \( \{ \tilde{C}_1^{CF}, \ldots, \tilde{C}_O^{CF} \} \) the market-clearing utility values that clear the market using \( \theta_{g}^{CF} = \tilde{\theta}_g \), and denote by \( \{ C_1^{CF}, \ldots, C_O^{CF} \} \) the market-clearing utility values that clear the market using \( \theta_{g}^{CF} = \tilde{\theta}_g \). Then \( \{ \tilde{C}_o^{CF} \} \) will satisfy \( \tilde{C}_o^{CF} = C_o^{CF} e^{\Delta_o} \forall o \) for some set of origin type-specific constants \( \{ \Delta_o \} \) that is invariant to the choice of \( f_{CF}(\*) \) and \( g_{CF}(\*) \).

Proof: I prove Proposition 1 by construction.

Let \( z(i, j, k) = 1(m(j) = m(k)) \) represent an indicator that takes on the value of 1 if the firms associated with positions \( j \) and \( k \) are the same, and 0 otherwise. Recall also that all worker transitions assigned to the same transition group \( g \) share values of the worker and establishment characteristics that define the worker’s origin and position’s destination types \( o \) and \( d \), respectively, as well as the value of the indicator \( z(i, j, k) \). Thus, one can write \( o(g), d(g) \) and \( z(g) \) for any group \( g \). Let the origin types be ordered (arbitrarily) from \( o = 1 \ldots o = O \), and let the destination types be ordered (arbitrarily) from \( d = 1 \ldots d = D \). Let \( g(o, d, z) \) denote the group associated with origin type \( o \), destination type \( d \), and existing worker indicator \( z \). Assume that the set \( \Theta_{D^{-in-D}} = \{ (\theta_{g'g}, \theta_{g''g''}) : (g, g', g'', g''') : o(g) = o(g''), o(g') = o(g'''), d(g') = d(g''), d(g'') = d(g''') \} \) is known, since a consistent estimator for each element of the set can be obtained via adjusted log odds ratios, as described in Section 3.

Consider defining the following set of alternative group-level joint surplus values \( \tilde{\Theta} = \{ \tilde{\theta}_g \} \) as follows:

\[
\tilde{\theta}_{g'} = 0 \forall g' : (o(g') = 1 \text{ and/or } d(g') = 1) \text{ and } z(g') = 0
\]

\[
\tilde{\theta}_{g'} = \frac{\theta_{g'g} - \theta_{g(1,d(g'),0)}}{\sigma} \forall g' : (d(g') \neq 1 \text{ and/or } z(g') \neq 0)
\]

(40)
Under the definitions in (40) and (41), we have:

\[
\frac{(\tilde{\theta}_g - \tilde{\theta}_{g'}) - (\tilde{\theta}_{g''} - \tilde{\theta}_{g'''})}{\sigma} = \frac{(\theta_g - \theta_{g'}) - (\theta_{g''} - \theta_{g'''})}{\sigma}
\]

\forall (g, g', g'', g''') : o(g) = o(g''), o(g') = o(g'''), d(g) = d(g'), d(g'') = d(g''') \tag{42}

Thus, the appropriate difference-in-differences using elements of \(\tilde{\Theta}\) match the corresponding difference-in-differences of true surpluses in \(\Theta\), so that all of the information about \(\Theta\) contained in the identified set \(\Theta - D\) is retained. Furthermore, unlike the true set \(\Theta\), the construction of \(\tilde{\Theta}\) only requires knowledge of \(\Theta - D\).

Next, note that the elements of \(\tilde{\Theta}\) can be written in the following form:

\[
\tilde{\theta}_g = \theta_g + \Delta^1_{o(g)} + \Delta^2_{d(g)} \quad \forall g \in \mathcal{G}, \text{ where}
\]

\[
\Delta^1_{o(g)} = \theta_{g(o(g),1,0)} - \theta_{g(1,1,0)} \tag{44}
\]

\[
\Delta^2_{d(g)} = \theta_{g(1,d(g),0)} \tag{45}
\]

where \(\mathcal{G}\) is the set of all possible transition groups. In other words, each alternative surplus \(\tilde{\theta}_g\) equals the true surplus \(\theta_g\) plus a constant \((\Delta^1_{o(g)})\) that is common to all groups featuring the same origin type and a constant \((\Delta^2_{d(g)})\) that is common to all groups featuring the same destination type.

Next, recall that there exists a unique aggregate assignment associated with each combination of marginal origin and destination type distributions \(f^{CF}(o)\) and \(h^{CF}(d)\) and set of group-level surpluses, including \(\tilde{\Theta}\). Let \(\tilde{P}^{CF}(\ast) \equiv P^{CF}(\ast|\tilde{\Theta}, \tilde{C}^{CF}_1, \ldots, \tilde{C}^{CF}_O)\) represent the unique counterfactual assignment that results from combining arbitrary marginal distributions \(f^{CF}(o)\) and \(h^{CF}(d)\) with the set \(\tilde{\Theta}\). \(\tilde{C}^{CF} = [1, \tilde{C}^{CF}_2 \ldots \tilde{C}^{CF}_O]\) denotes the vector of mean exponentiated utility values for each origin type \(o\) (with \(\tilde{C}^{CF}_1\) normalized to 1) that solves the system of excess demand equations below, and thus yields \(\tilde{P}^{CF}(g) \forall g \in \mathcal{G}\) when plugged into equation (23) along with the elements of \(\tilde{\Theta}\):

\[
\sum_{d \in D} h^{CF}(d) \left( \sum_{g : o(g) = 2} P^{CF}(g|d, \tilde{\Theta}, \tilde{C}^{CF}) \right) = f^{CF}(2) \tag{46}
\]

I wish to show that \(\tilde{P}^{CF}(\ast) \equiv P^{CF}(\ast|\tilde{\Theta}, \tilde{C}^{CF})\) will be identical to the alternative unique counterfactual equilibrium assignment \(P^{CF}(\ast|\Theta, C^{CF})\) that combines the same arbitrary marginal distributions \(f^{CF}(o)\) and \(h^{CF}(d)\) with the set \(\Theta\) instead of \(\tilde{\Theta}\). Here, \(C^{CF} = [1, C^{CF}_2 \ldots C^{CF}_O]\) denotes a vector of \(o\)-type-specific mean exponentiated utility values that
clears the market by satisfying the following alternative excess demand equations:\textsuperscript{44}

\[
\sum_{d \in D} h_{CF}^C(d)(\sum_{g: o(g) = 2} P_{CF}^C(g|d, \Theta, \tilde{C}^C))^e = f^C(2) \\
\vdots \\
\sum_{d \in D} h_{CF}^C(d)(\sum_{g: o(g) = O} P_{CF}^C(g|d, \Theta, \tilde{C}^C))^e = f^C(O) 
\]

Since all other terms are unchanged between the systems (46) and (47), it suffices to show that \(P^C_{CF}(g|d, \tilde{\Theta}, \tilde{C}^C) = P^C_{CF}(g|d, \Theta, C^C)\) \(\forall g \in G\) for some vector \(C^C\). Consider the following proposed vector \(C^C\):

\[
C_o^C = \tilde{C}_o^C e^{\frac{\Delta^1_o}{\sigma}} \forall o \in [2, \ldots, O] 
\]

where \(\Delta^1_o\) is as defined in (44). For an arbitrary choice of \(g\), we obtain:

\[
P^C_{CF}(g|d(g), \tilde{\Theta}, \tilde{C}^C) \\
= \frac{\tilde{\sigma}_{g}^{C^C}}{\tilde{\sigma}_{g|o(g), d(g)}^{C^C} f_{CF}^{C^C}(o(g)) \tilde{C}_o^C} \\
\sum_{o' \in O} \sum_{g' \in (a, d)} \frac{\tilde{\sigma}_{g'}^{C^C}}{\tilde{\sigma}_{g'|o(g')}, d(g)} f_{CF}^{C^C}(o') \tilde{C}_o^C \\
= \frac{\tilde{\sigma}_{g}^{C^C}}{\tilde{\sigma}_{g|o(g), d(g)}^{C^C} f_{CF}^{C^C}(o(g)) C_o^C} \\
\sum_{o' \in O} \sum_{g' \in (a, d)} \frac{\tilde{\sigma}_{g'}^{C^C}}{\tilde{\sigma}_{g'|o(g')}, d(g)} f_{CF}^{C^C}(o') C_o^C \\
= \frac{\tilde{\sigma}_{g}^{C^C}}{\tilde{\sigma}_{g|o(g), d(g)}^{C^C} f_{CF}^{C^C}(o(g)) C_o^C} \\
\sum_{o' \in O} \sum_{g' \in (a, d)} \frac{\tilde{\sigma}_{g'}^{C^C}}{\tilde{\sigma}_{g'|o(g')}, d(g)} f_{CF}^{C^C}(o') C_o^C \\
= P^C_{CF}(g|d(g), \Theta, C^C) 
\]

This proves that \(P^C_{CF}(g|d, \Theta, C^C)\) also satisfies the market clearing conditions (47) above, and will therefore be the unique group-level assignment consistent with marketwide equilibrium and stability. Thus, I have shown that the counterfactual assignment that is recovered when using an alternative set of surpluses \(\tilde{\Theta}\) derived from the identified set \(\Theta^{D_{-in} - D}\) will in fact equal the counterfactual assignment I desire, which is based on the true set of joint surplus values \(\Theta\). Furthermore, while origin-type specific mean utility values \(\tilde{C}^C\) that clears the market given \(\tilde{\Theta}\) will differ for each origin type from the corresponding vector

\textsuperscript{44}Note that I have suppressed the dependence of \(P^C_{CF}(g|\Theta, C^C, f_{CF}^C(a), h_{CF}^C(d), \tilde{\sigma}_{g|o, d})\) on \(f_{CF}^C(a), h_{CF}^C(d), \text{ and } \tilde{\sigma}_{g|o, d}\) because these are held fixed across the two alternative counterfactual simulations.
based on the true surplus set $\Theta$, these differences are invariant to the marginal origin and destination distributions $f^{CF}(o)$ and $h^{CF}(d)$ used to define the counterfactual. This implies that differences in utility gains caused by alternative counterfactuals among origin groups are identified, permitting comparisons of the utility incidence of alternative labor supply or demand shocks. This concludes the proof.

A3 Estimating the Value of $\sigma$

I attempt to estimate $\sigma$, the standard deviation of the unobserved match-level component $\epsilon_{ij(i)k}$, by exploiting the fact that the composition of U.S. origin and destination job matches $f^y(o)$ and $h^y(d)$ evolved across years $y$. Specifically, I estimate the set of group-level surpluses $\{\theta^g_{2007}\}$ from the observed 2007-2008 matching. Then, holding these surplus values fixed, I combine $\{\theta^g_{2007}\}$ with $f^y(o)$ and $h^y(d)$ from each other year $y \in [1993, 2010]$ to generate counterfactual assignments and changes in scaled mean (exponentiated) utility values $\{C^{CF}_o\}$ for each origin type. These counterfactuals predict how mean worker utilities by skill/location combination could have been expected to evolve over the observed period given the observed compositional changes in labor supply and demand had the underlying surplus values $\{\theta^g\}$ been constant and equal to $\{\theta^g_{2007}\}$ throughout the period.

To the extent that most of evolution in the utility premia enjoyed by workers in particular locations and skill categories was due primarily to changes in supply and demand composition rather than changes in the moving costs, recruiting costs, tastes, and relative productivities that compose the joint surplus values $\{\theta^g\}$, these counterfactual predictions will be reasonable approximations of the realized evolution of ex post utility over time by origin type. Recall that $C^{CF}_o \approx \frac{1}{|o|} \sum_{i:o(i,j(i))=o} e^{\frac{r_{i}^{CF,y}}{\sigma y}}$. Thus, if ex post utility $r_{i}^{CF}$ does not vary too much across individuals within an origin type, so that Jensen’s inequality is near equality and $\frac{1}{|o|} \sum_{i:o(i,j(i))=o} e^{\frac{-r_{i}^{CF,y}}{\sigma y}} \approx e^{\frac{-r_{i}^{CF,y}}{\sigma y}}$, then taking logs yields $ln(C^{CF,y}_o) \approx \frac{r_{i}^{CF,y}}{\sigma y}$.

Next, I form the corresponding changes in observed annual earnings from origin to destination match for each origin type in each year, $\bar{E}arn^{y+1}_o - \bar{E}arn^y_o$. I then run the following regression at the $o$-type level for each year $y \in [1993 - 2011]$:

$$\bar{E}arn^{y+1}_o - \bar{E}arn^y_o = \beta^y_0 + \beta^y_1(ln(C^{CF,y+1}_o) - ln(C^{CF,y}_o)) + \nu^y_o \tag{50}$$

Recall that the $\nu^y_o$ values represent predicted money metric utility gains, and are thus denominated in dollars. However, even if the surplus values $\{\theta^g\}$ are time invariant over the chosen period (and the other assumptions of the assignment model specified above all hold, including the approximations just described), dollar-valued mean utility gains would not equal mean annual earnings gains for a given origin type if its workers systematically moved

---

Note that while worker earnings in origin job matches were used to assign workers to skill categories, to this point I have not used observed worker earnings in destination positions to identify any other parameters.
to jobs featuring better or worse amenities, avoided more moving/recruiting training costs \( c^l(j,k) \), or moved to jobs featuring better or worse continuation values. However, if such changes in other sources of utility nearly cancel out among workers assigned to the same origin type (for all origin types), then \( r^F_{\alpha} \) should approximately equal \( \bar{\text{Earn}}^{y+1}_{o} - \bar{\text{Earn}}^{y}_{o} \). This implies that \( \beta^y \approx \sigma^y \).

As noted in Section 5.2, the origin type space depends on which location is considered the target location for the shock, with the geographic units that partially define origin types becoming more aggregated farther from the shock. To address this issue, in practice I constructed separate true and counterfactual earnings changes and estimated 50 for the collapsed origin type spaces associated with each possible target PUMA among the sample states, and averaged the estimates of \( \beta \) across all regressions satisfying a minimum \( R^2 \) threshold of .1 to obtain \( \hat{\beta}^y \). The estimates of \( \hat{\beta}^y \) are fairly consistent across years, so I use the mean estimate across all years, \( \bar{\sigma} = 8,430 \), to produce dollar values for all the results relating to utility gains presented in the paper.

Clearly, given the additional strong assumptions required, this approach represents a relatively crude attempt to calibrate \( \sigma \). Indeed, further efforts could conceivably be taken to exclude origin types \( o' \) whose surplus values \( \{ \theta_g : o(g) = o' \} \) were known to be changing over the chosen time period, or to allow \( \theta_g \) to evolve in a particular parametric fashion.\(^47\) In fact, Galichon and Salanié (2015) discuss how a vector of \( \sigma \) values associated with different types or combinations of types based on observed characteristics might potentially be jointly estimated with other model parameters (thereby allowing heteroskedasticity across types in the idiosyncratic match component). Since our focus is primarily on examining relative incidence across different origin types from shocks featuring different changes in labor demand composition, I opted for the simpler, more transparent approach.

### A4 Using Transfers to Decompose the Joint Surpluses \( \{ \theta_g \} \)

This appendix examines whether observing equilibrium transfers, denoted \( w_{ik} \), allows the identification of additional parameters of interest. In Choo and Siow (2006)'s assignment model, the unobserved match-level heterogeneity is assumed to take the form \( \epsilon_{ijk} = \epsilon^1_{o(i,j)k} + \epsilon^2_{ijd(k)} \), so that aggregate surplus is left unchanged when two pairs of job matches \( (i,k) \)

\(^{46}\)A few PUMAs and states experienced relatively little year-to-year change in the distribution of employment across destination types, so that the counterfactual earnings forecasts predicted true earnings changes poorly. In this case, the \( R^2 \) from the regression was very low, \( \hat{\beta}^y \) was badly identified. The results become far more stable across the remaining alternative type spaces when a minimum \( R^2 \) was imposed to eliminate the few badly identified estimates, which tended to produce outliers.

\(^{47}\)In the actual implementation, I do allow the set of \( \theta_g \) used to generate the counterfactual prediction to evolve over time in an extremely restricted fashion: I allow the relative payoff of retaining existing workers relative to hiring new workers to evolve over time to match the share of workers who stay at their dominant jobs in each observed year. I do this because the well-chronicled decline in job-to-job mobility during this time period is strongly at odds with the assumption that \( \theta_g \) is completely time invariant. 

---

85
and \((i', k')\) belonging to the same group \(g\) swap partners. The elimination of any true \((i, k)\) match-level surplus component implies that equilibrium transfers cannot vary among job matches belong to the same group \(g\), so that \(w_{ik} = w_g(i, k) \forall (i, k)\).\footnote{If \(w_{ik} > w_{i'k'}\) for any two matched pairs \((i, k)\) and \((i', k')\) such that \(g(i, k) = g(i', k')\), then \((i', k')\) would form a blocking pair by proposing a surplus split between them featuring a transfer between \(w_{ik}\) and \(w_{i'k'}\), thus undermining the stability of the proposed matching.} Galichon and Salanié (2015) show that under this assumption about the form of unobserved heterogeneity, observing the (common) group-level transfers \(w_g\) would be sufficient to decompose the group-level mean joint surplus \(\theta_g\) into the worker and position’s respective pre-transfer payoffs, which I denote \(\theta^f_g\) and \(\theta^l_g\), respectively.

Because the model proposed in section 2.3 does not impose the additive separability assumption \(\epsilon_{ijk} = \epsilon_{o(i,j)k} + \epsilon_{ijd(k)}\), equilibrium transfers will in general vary among \((i, k)\) pairs belonging to the same group \(g\). Indeed, given that I observe substantial earnings variance within observed groups \(g\) regardless of the worker, position, and job transition characteristics used to define \(g\), the Choo and Siow (2006) restriction on the nature of unobserved match-level heterogeneity would be strongly rejected in the labor market context.

However, one can still consider the value for identification of the observed transfers \(\{w_{ik}\}\). Recall from section 2.2 that equilibrium transfers are related to equilibrium worker and position payoffs via:

\[
\begin{align*}
w_{ik} &= \pi^f_{ik} - q_k \\
\pi^l_{ik} &= r_i - \pi^f_{ik}
\end{align*}
\]

Next, recall from equation (20) that under Assumptions 1 and 2 the log odds that a randomly chosen position from arbitrary destination type \(d\) will choose a worker whose hire would be assigned to group \(g_1\) relative to \(g_2\) are given by:

\[
\ln\left(\frac{P(g_1|d)}{P(g_2|d)}\right) = \ln(P(g_1|d)) - \ln(P(g_2|d)) = \frac{\theta_{g_1}}{\sigma} + \ln(\overline{S}_{g_1|o(g_1),d}) + \ln(f(o(g_1))) + \ln(C_{o(g_1)}) - \frac{\theta_{g_2}}{\sigma} - \ln(\overline{S}_{g_2|o(g_2),d}) - \ln(f(o(g_2))) - \ln(C_{o(g_2)})
\]

(53)

Since \(\ln(\overline{S}_{g_1|o(g_1),d}), \ln(\overline{S}_{g_2|o(g_2),d}), \ln(f(o(g_1)))\), and \(\ln(f(o(g_2)))\) are all observed (or, if a large sample is taken, extremely precisely estimated), one can instead form adjusted log odds as in (21):

\[
\ln\left(\frac{\hat{P}_{g_1|d}/(\overline{S}_{g_1|o(g_1),d})f(o(g_1))}{\hat{P}_{g_2|d}/(\overline{S}_{g_2|o(g_2),d})f(o(g_2))}\right) = \left(\frac{\theta_{g_1} - \theta_{g_2}}{\sigma}\right) + (\ln(C_{o(g_1)}) - \ln(C_{o(g_2)}))
\]

(54)

Under Assumption 1, \(C_o\) is the mean of exponentiated (and rescaled) equilibrium utility.
payoffs owed to workers \( i : o(i) = o \):

\[
C_o = \frac{1}{|o|} \sum_{i:o(i,j(i))=o(g)} e^{-\frac{r_i}{\sigma}} \approx \sum_{1_k} \sum_{i:g(i,i,k)=g} e^{-\frac{r_i}{\sigma}} \forall k
\] (55)

Plugging (52) into (55) and then (55) into (54) yields:

\[
\ln\left(\frac{\hat{P}_{g1|d/(S_{g1}|o(g1),df(o(g1)))}}{\hat{P}_{g2|d/(S_{g2}|o(g2),df(o(g2)))}\right)
= \left(\frac{\theta_{g1} - \theta_{g2}}{\sigma}\right) + \left(\ln\left(\frac{1}{|o|} \sum_{i:o(i,j(i))=o(g1)} e^{-\frac{w_{ik} + \pi_{lk}^1}{\sigma}}\right) - \ln\left(\frac{1}{|o|} \sum_{i:o(i,j(i))=o(g2)} e^{-\frac{w_{ik} + \pi_{lk}^2}{\sigma}}\right)\right)
\] (56)

It is not immediately obvious how to use equation (58) to recover parameters of interest. Only when one adds further assumptions that are at odds with the structure of the model can one recover an expression that mirrors the one in Choo and Siow (2006). Specifically, suppose the following assumptions hold:

\[
\begin{align*}
    r_i &\approx r_{o(i)} \forall i : o(i, j(i)) = o \forall o \in O \\
    \pi_{lk}^1 &\equiv \pi_{g(i,k)}^1 \forall (i, k) : g(i,k) = g \forall g \in G \\
    w_{ik} &\equiv w_{g(i,k)} \forall (i, k) : g(i,k) = g \forall g \in G
\end{align*}
\] (57)

I suspect that these assumptions will are extremely unlikely to hold in any stable matching if there is meaningful variance in the unobserved match surplus component \( e_{ij(i)k} \) among \((i,k)\) pairs assigned to the same group \( g \), whose extreme value distribution is the basis for the logit closed-forms used for conditional choice probabilities above. Nonetheless, they yield:

\[
\ln\left(\frac{\hat{P}_{g1|d/(S_{g1}|o(g1),df(o(g1)))}}{\hat{P}_{g2|d/(S_{g2}|o(g2),df(o(g2)))}\right)
= \left(\frac{\theta_{g1} - \theta_{g2}}{\sigma}\right) + \left(\ln\left(\frac{1}{|o|} \sum_{i:o(i,j(i))=o(g1)} e^{-\frac{w_{ik} + \pi_{lk}^1}{\sigma}}\right) - \ln\left(\frac{1}{|o|} \sum_{i:o(i,j(i))=o(g2)} e^{-\frac{w_{ik} + \pi_{lk}^2}{\sigma}}\right)\right)
\] (58)

Given an estimate of \( \sigma \) based on multiple markets (as described in Appendix A3) and data on mean annual earnings for each transition group \( g \in G \), one could identify the difference in the position component of the joint surplus for arbitrary groups \( g_1 \) and \( g_2 \). This provides information about the relative profit contributions of different types of workers for each
type of position before such workers salaries are considered. Note that one could still not separate the training cost, recruiting cost, current revenue contribution, and continuation value components of $\theta_g^f$ without additional data.

A similar progression using adjusted log odds based on the worker side conditional probabilities $P(g_1|o_1)$ and $P(g_2|o_1)$ would yield an estimate of the corresponding difference in the worker components of the joint surplus $\theta_{g_1}^l - \theta_{g_2}^l$ for any two groups featuring the same origin worker type. Since one such group could represent nonemployment, this approach would provide estimates of the desirability of working at various types of firms in various locations for zero pay relative to nonemployment. These values identify the reservation salary necessary to convince each origin worker type to take (or continue) a position of each destination type. As with firms, one could not disentangle the moving cost, search cost, non-wage amenity value, and continuation value components of the surplus without further data.

Because 1) I deem the assumptions (57) to be antithetical to the spirit of the model and at odds with the data, and 2) other than estimating $\sigma$, the use of transfers is not necessary to fulfill the primary aim of the paper, evaluating the utility and profit incidence across worker and position types of alternative local labor demand shocks, I do not make further use of the observed annual earnings distributions in the destination period $y$ in any aggregate labor market transition $(y-1, y)$ in this paper.

A5 Removing Spurious U-to-U, E-to-U, and U-to-E Transitions

The inability to observe workers working in states that did not approve the use of their LEHD data for my project introduces the possibility that many of the workers who are not observed working in a given year in my sample (despite being observed working in other years) are in fact working in out-of-sample state. This appendix describes how supplementary data from the harmonized American Community Survey (hereafter ACS) series created by IPUMS along with official unemployment statistics from the Bureau of Labor Statistics (hereafter BLS) were used to mitigate this problem.

Since distinct procedures were used to impute the number of unemployment-to-employment transitions (denoted U-to-E) and employment-to-unemployment transitions (E-to-U) is distinct from the procedure used to impute unemployment-to-unemployment transitions (U-to-U), I discuss the two procedures separately.
A5.1 Unemployment-to-Employment and Employment-to-Unemployment Transitions

Note first that a transition count must be generated for each transition group $g$ classified as a U-to-E transition, which consists of a combination of origin location, age group ($< 25$ or $\geq 25$), destination location, and establishment size quartile, average pay quartile, and industry supersector. Because the ACS does not collect data on establishment size or average pay, and because the 1% ACS sample is too small to generate accurate counts at the tract-to-tract or even PUMA-to-PUMA transition level, I begin by pooling all ACS years between 2005 and 2016 to create counts of U-to-E transitions by combination of origin state, age group, destination state, and destination industry supersector.\footnote{While pooling years rather than generating year-specific estimates of the share of U-to-E transitions that are spurious in my LEHD sample may generate some measurement error, the state-by-age-by-state-by-industry group space is sufficiently fine that the sampling error in group-level counts from any one year can be substantial, and pooling across years alleviates this problem. Importantly, the ACS does not distinguish between unemployed and out-of-labor force workers in the origin year, so all initially nonemployed workers who take a job in the destination year are considered in the labor force from the perspective of the destination year job search.} I then create counts of E-to-E transitions for the same aggregated transition groups for the subset of origin locations that are outside the available sample of states. Next, for each state/age/state/industry combo, I divide the count of U-to-E transitions featuring a within-sample origin state and a within-sample destination state (true in-sample U-to-E transitions) by the count of all transitions that might be construed as U-to-E transitions in my LEHD sample, whether true or spurious. The spurious transitions consist of U-to-E transitions and E-to-E transitions featuring an out-of-sample origin state and a within-sample destination state. This set of ratios estimates the expected share of LEHD U-to-E transitions that are not spurious for each state/age/state/industry combo (i.e. $P(\text{True } U \to E|\text{loc}(o), \text{age}(o), \text{loc}(d), \text{industry}(d))$

I then multiply the LEHD count for each transition group associated with U-to-E transitions by the estimate $P(\text{True } U \to E|\text{loc}(o), \text{age}(o), \text{loc}(d), \text{industry}(d))$ for the appropriate state/age/state/industry combo of the chosen group $g$. This re-scales all such counts so that they match the estimated true $U \to E$ counts for each state/age/state/industry combo.

The procedure for imputing E-to-U transitions is roughly analogous. The same 2005-2016 ACS data is used to create pooled counts of aggregated groups corresponding to E-to-U transitions, this time defined by origin state, origin earnings quartile, origin age group, and destination state. Since all destination locations are eventually treated as a single “nonemployment” location by the two-sided model used to generate the labor demand shock simulations, the destination state is only used to distinguish E-to-U transitions in which the worker moves to an out-of-sample state to search for a job (to be excluded from the sample) from those in which the worker moves to an in-sample state. A corresponding count is generated of E-to-E transitions at the same aggregated group level for which the origin
state is in the sample but destination state is out-of-sample, as well as E-to-NE transitions regardless of destination state. Both of these constitute additional sources of spurious E-to-U transitions in the LEHD. A ratio is computed for each origin state/earnings/age combo of the “true” E-to-U transition counts divided by the “spurious” E-to-U transitions. These ratios estimate the expected share of LEHD E-to-U transitions that are not spurious for each origin state/earnings/age combo \( P(\text{True}\ E \to U | \text{loc}(o), \text{earn}(o), \text{age}(o)) \).

I then multiply the LEHD count for each transition group associated with E-to-U transitions by the estimate \( P(\text{True}\ E \to U | \text{loc}(o), \text{earn}(o), \text{age}(o)) \) for the appropriate state/earnings/age combo of the chosen group \( g \). This re-scales all such counts so that they match the estimated true \( E \to U \) counts for each origin state/earnings/age combination.

### A5.2 Unemployment-to-Unemployment Transitions

Since U-to-U transitions are particularly prone to mismeasurement in the LEHD, I rely particularly heavily on ACS and BLS data to generate these counts. Specifically, I begin by creating counts of NE-to-U transitions in the ACS for each origin state/age category combination for each year. Because these include both OLF-to-U and U-to-U transitions, I then rescale these counts by multiplying by \( \frac{\# \text{ U-to-U BLS}}{\# \text{ NE-to-U ACS}} \). Here, \( \# \text{ NE-to-U ACS} \) is the sum of NE-to-U transition counts across all state-age combos for the chosen year, and \( \# \text{ U-to-U BLS} \) is the average across the final three months of the year of the BLS reported count of workers who have been unemployed more than 52 weeks. This rescaling ensures that the total number of imputed U-to-U transitions will match the BLS long-term unemployment count.

For years prior to 2001, for which ACS NE-to-U counts cannot be constructed, I take the 2001 NE-to-U counts for each origin state/age category combination and multiply them by the ratio of the BLS long-term unemployment counts in the chosen year and 2001, respectively, so that the sum of these imputed counts will at least match the BLS long-term unemployment count in the chosen year.\(^50\)

Finally, since the U-to-U transition groups used in the two-sided matching model in some cases feature census tract or PUMA as the origin location category, I use the conditional distribution of origin tract conditional on origin state among NE-to-NE transitions in the LEHD (some of which may be spurious) to distribute the U-to-U counts that were originally computed at the state-age level across tracts and PUMAs.

\(^{50}\)Because I use a 50% random sample of all LEHD transitions in this version, I multiply these estimated U-to-U counts by .5.
A6 Smoothing Procedure

In this appendix I describe how I smooth the empirical distribution of transitions across transition groups, $\hat{P}(g)$, prior to estimation in order to generate accurate estimates of the elements of the identified set of joint surplus difference-in-differences $\Theta_{D-in-D}$. I smooth for two reasons. First, such smoothing serves as a “noise infusion” technique that removes the risk that the identity of any particular individual or establishment could be revealed by any of the estimates presented in the paper, as required of any research results generated from confidential microdata in Federal Statistical Research Data Centers (FSRDCs). Second, smoothing is necessary because there are sufficiently few observations per transition group that many transition groups are either rarely (or never) observed in a given cross-sectional transition despite substantial underlying matching surpluses simply due to sampling error. Essentially, $\hat{P}(g)$ is only a consistent estimator of $P(g)$ as the number of observed worker transitions per group $I/G$ approaches infinity.

I overcome this sampling error problem by assuming that the underlying frequency $P(g)$ with which a randomly chosen transition belongs to a particular transition group is a smooth function of the observed characteristics that define group $g$. This permits the use of a kernel density estimator that computes a weighted average of the empirical probabilities $\hat{P}(g')$ of “nearby” groups $g'$ that feature “similar” vectors of characteristics to generate a well-behaved approximation of $P(g)$ from the noisy empirical distribution $\hat{P}(g)$.

Such smoothing introduces two additional challenges. First, excessive smoothing across other transition groups erodes the signal contained in the data about the degree of heterogeneity in the relative surplus from job transitions featuring different combinations of worker characteristics, establishment characteristics, and origin and destination locations. Since highlighting the role of such heterogeneity in forecasting the incidence of labor market shocks is a primary goal of the paper, decisions about the appropriate bandwidth must be made with considerable thought. The second, related challenge consists of identifying which of the worker and position characteristics that defines other groups makes them “similar”, in the sense that the surplus $\{\theta_{g'}\}$ is likely to closely approximate the surplus whose estimate I wish to make more precise $\theta_g$.

Recall that each group $g \equiv g(o,d,z)$ is a combination of 1) the origin establishment location (which I denote $loc(o)$) and workers’ initial earnings quartile (or nonemployment status) at the origin establishment (denoted $earn(o)$); 2) the destination establishment’s location ($loc(d)$), establishment size category ($f_size(d)$), establishment average earnings category ($f_earn(d)$), and industry supersector ($ind(d)$); and 3) the indicator $z(i,j,k)$ for whether establishment $j$ and establishment $k$ are the same, so that worker $i$ is a job stayer rather than a mover (denoted $stayer(g)$).

Given our goal of accurate characterizing incidence at a very low level of geographic aggregation, I wish to preserve as accurately as possible any signal in the data about the
structure of spatial ties between nearby local areas. Thus, wherever possible our kernel estimator should place non-zero weight only on alternative groups \( g' \) that share the same origin and destination locations (\( \text{loc}(o(g)) = \text{loc}(o(g')) \) and \( \text{loc}(d(g)) = \text{loc}(d(g')) \)). Similarly, I suspect that the combination of the non-location characteristics establishment size, establishment average worker earnings, and firm industry is likely to be more important than location in determining the skill category of worker (proxied by initial earnings quartile) that generates the most surplus. To develop a smoothing approach that embodies these principles, I exploit the fact that \( P(g) \) can be decomposed via:

\[
P(g) = P(g|d(g))h(d(g)) = P([o(g), d(g), z(g)]|d)h(d(g))
\]

\[
= P([\text{loc}(o(g)), \text{earn}(o(g)), \text{stayer}(g)]|d)h(d(g))
\]

\[
= P(\text{loc}(o(g))|\text{earn}(o(g)), \text{stayer}(g), d)P([\text{earn}(o(g)), \text{stayer}(g)]|d)h(d(g))
\]

\[
= 1(\text{stayer}(g) = 1)P(\text{loc}(o(g))|\text{earn}(o(g)), 1(\text{stayer}(g) = 1), d)P([\text{earn}(o(g)), 1(\text{stayer}(g) = 1)]|d)h(d(g))
\]

\[
+ 1(\text{stayer}(g) = 0)P(\text{loc}(o(g))|\text{earn}(o(g)), 1(\text{stayer}(g) = 0), d)P([\text{earn}(o(g)), 1(\text{stayer}(g) = 0)]|d)h(d(g))
\]

\[
= 1(\text{stayer}(g) = 1)1(\text{loc}(o(g)) = \text{loc}(d(g)))P([\text{earn}(o(g)), 1(\text{stayer}(g) = 1)]|d)h(d(g))
\]

\[
+ 1(\text{stayer}(g) = 0)P(\text{loc}(o(g))|\text{earn}(o(g)), 1(\text{stayer}(g) = 0), d)P([\text{earn}(o(g)), 1(\text{stayer}(g) = 0)]|d)h(d(g))
\]

\[(59)\]

where the first two lines use the law of total probability and the set of characteristics that define \( o(g) \) and \( z(g) \), the third line uses the fact that the \( z(g) \equiv \text{Stayer}(g) \) only takes on two values (0 for job movers and 1 for job stayers), and the last line uses the fact that \( P(\text{loc}(o(g))|\text{earn}(o(g)), 1(\text{stayer}(g) = 1), d) = 1(\text{loc}(o(g)) = \text{loc}(d(g))) \), since a potential stayer associated with a particular destination type must have already been working at the same location in the origin period (by virtue of being a job stayer, since I treat firms that switch locations as different firms for computational reasons).

I use separate kernel density estimator procedures to estimate each of \( P(\text{loc}(o(g))|\text{earn}(o(g)), 1(\text{stayer}(g) = 0), d(g)) \), \( P(\text{earn}(o(g)), 1(\text{stayer}(g) = 0)|d(g)) \), and \( P(\text{earn}(o(g)), 1(\text{stayer}(g) = 1)|d(g)) \).

Consider first the estimation of \( P(\text{loc}(o(g))|\text{earn}(o(g)), 1(\text{stayer}(g) = 0), d(g)) \), the conditional probability that a particular new hire would be originally located at location \( \text{loc}(o) \), given the hired worker’s initial earnings category and the destination type \( d \) of the hiring position. Let \( K^{\text{dist}}(g, g') \) represent the metric capturing how similar an alternative group \( g' \) is to \( g \) for the purpose of estimating the propensity for firms of type \( d \) to hire workers from a particular location (conditional on skill level). As discussed above, wherever possible I only assign non-infinite distance \( K^{\text{dist}}(g, g') < \infty \) (which corresponds to non-zero weight) to empirical conditional probabilities \( P(\text{loc}(o(g'))|\text{earn}(o(g')), 1(\text{stayer}(g') = 0), d(g')) \) of alternative groups \( g' \) that feature both the same origin location \( \text{loc}(o(g')) = \text{loc}(o(g)) \) and
the same destination location \( loc(d(g')) = loc(d(g)) \).\(^{51}\)

\( K^{\text{dist}}(g, g') \) assigns the smallest distance to alternative groups \( g' \) that also feature the same destination type \( (d(g') = d(g)) \), so that \( g \) and \( g' \) only differ in the initial earnings category of hired workers. The closer \( \text{earn}(o(g')) \) is to \( \text{earn}(o(g')) \), the smaller is the assigned distance \( K^{\text{dist}}(g, g') \), but the profile flattens so that all groups \( g' \) that differ from \( g' \) only due to \( \text{earn}(o(g')) \) contribute to the weighted average. \( K^{\text{dist}}(g, g') \) assigns larger (but still noninfinite) distance to groups \( g' \) featuring destination types that also differ on establishment size, establishment avg. earnings, or industry dimensions. The more different the establishment composition of the group, the smaller is its weight, with the profile again flattening so that all groups \( g' \) featuring the same origin and destination locations receive non-zero weight. Thus, groups with less similar worker and establishment characteristics receive non-negligible weight only when there are too few observations from groups featuring more similar worker and establishment characteristics to form reliable estimates. The weight assigned to a particular alternative group \( g' \) also depends on the number of observed new hires made by \( d(g') \) at a particular skill level \( \text{earn}(o(g')) \), denoted \( N^{\text{dist}}(g') \) below, since this determines the signal strength of the empirical conditional choice probability \( P(loc(o(g')|\text{earn}(o(g'))), 1(\text{stayer}(g') = 0), d(g')) \). Thus, we have:

\[
P(loc(o(g)|\text{earn}(o(g)), 1(\text{stayer}(g) = 0), d(g)) \approx \sum_{g'} \frac{\phi(K^{\text{dist}}(g', g)N^{\text{dist}}(g'))}{\sum_{g''} \phi(K^{\text{dist}}(g'', g)N^{\text{dist}}(g''))} P(loc(o(g')|\text{earn}(o(g'))), 1(\text{stayer}(g') = 0), d(g')) \tag{60}
\]

where \( \phi(\cdot) \) is the normal density function (used as the kernel density), and represents the weight given to a particular nearby transition group \( g' \).

Next, consider the estimation of \( P(\text{earn}(o(g)), 1(\text{stayer}(g) = 1)|d) \) and \( P(\text{earn}(o(g)), 1(\text{stayer}(g) = 0)|d') \), the conditional probabilities that either a job stayer or mover originally paid at a particular earnings quartile (or possibly non-employed for movers) will be hired to fill a position of destination type \( d \). Let \( K^{\text{earn/move}}(g, g') \) and \( K^{\text{earn/stay}}(g, g') \) represent the metrics capturing how similar alternative groups \( g' \) are to \( g \) for the purpose of estimating the propensity for firms of type \( d \) to hire (or retain) workers at particular skill levels.

\( K^{\text{earn/move}}(g, g') \) and \( K^{\text{earn/stay}}(g, g') \) each assign infinite distance (translating to zero weight) to groups \( g' \) featuring different combinations of establishment size, average worker earnings, or industry than the target group \( g \). \( K^{\text{earn/move}}(g, g') \) \( (K^{\text{earn/stay}}(g, g')) \) assigns small distances to the conditional probabilities associated with groups \( g' \) representing hiring new (retaining) workers from the same initial earnings (or nonemployment) category \( \text{earn}(o(g)) = \text{earn}(o(g')) \) among firms from the same destination type \( d(g) = d(g') \) but who are hiring workers from nearby locations. The distance metric increases in the tract

\(^{51}\)There are a very small number of destination and origin types are never observed in any transition. By necessity, I put positive weight on groups featuring nearby origin or destination locations in such cases.
pathlength between \( \text{loc}(o(g')) \) and \( \text{loc}(o(g')) \), but flattens beyond a threshold distance, so that groups featuring all origin worker locations (but same other characteristics) contribute to the estimate.

Larger (but finite) distance values for \( K_{\text{earn/move}}(g,g') \) and \( K_{\text{earn/stay}}(g,g') \) are assigned to conditional probabilities from groups \( g' \) that feature different (but nearby) destination locations (so \( d(g) \neq d(g') \) but the same combination of establishment size quartile, establishment average worker earnings quartile, and industry supersector. Again, the distance metric increases in the pathlength between \( \text{loc}(d(g)) \) and \( \text{loc}(d(g')) \), but eventually flattens at a large but non-infinite value. As before, the weight given to a group \( g' \) also depends on the precision of its corresponding number of total hires made by firms of the destination type \( d(g') \), which is proportional to \( h(d(g')) \).

Again, the motivation here is that targeted skill level and the decision to retain workers vs. hire new workers (conditional on the utilities bids required by workers in different locations) is likely to be driven to a greater extent by the type of production process (as proxied by size, mean worker earnings, and industry) than by the location of the establishment. Nonetheless, since I still suspect that there is unobserved heterogeneity in production processes conditional on our other establishment observables that might be spatially correlated, I place greater weight on the skill/retention decisions of geographically proximate firms. More distant firms receive non-negligible weight only when there are too few local observations to form reliable estimates. The estimators for \( P(\text{earn}(o(g)), 1(\text{stayer}(g) == 1)|d) \) and \( P(\text{earn}(o(g)), 1(\text{stayer}(g) == 0)|d) \) can thus be represented via:

\[
P(\text{earn}(o(g)), 1(\text{stayer}(g) == 0)|d) \approx \frac{\sum_{g'} \phi(K_{\text{earn/move}}(g',g) h(d(g')))}{\sum_{g''} \phi(K_{\text{earn/move}}(g'',g) h(d(g'')))} \hat{P}(\text{earn}(o(g'), 1(\text{stayer}(g') = 0)|d(g'))) \tag{61}
\]

\[
P(\text{earn}(o(g)), 1(\text{stayer}(g) == 1)|d) \approx \frac{\sum_{g'} \phi(K_{\text{earn/stay}}(g',g) h(d(g')))}{\sum_{g''} \phi(K_{\text{earn/stay}}(g'',g) h(d(g'')))} \hat{P}(\text{earn}(o(g'), 1(\text{stayer}(g') = 1)|d(g'))) \tag{62}
\]

Bringing the pieces together, this customized smoothing procedure has a number of desirable properties. First, by requiring the same origin and destination locations as a necessary condition for non-zero weight when estimating the propensity for particular destination types to hire workers from each location, one can generate considerable precision in estimated conditional choice probabilities without imposing any assumption about the spatial links between locations. Second, at the same, one can still use information contained in the hiring and retention choices of more distant firms to learn about the propensity for firms of different size, pay level, and industry to retain and hire workers at different skill levels and from nonemployment. Third, the procedure places non-trivial weight on transition groups featuring less similar worker and establishment characteristics only when there
are too few observed hires/retentions made by firms associated with groups featuring very similar characteristics to yield reliable estimates. Fourth, overall the estimated probabilities \( P(q|d) \) place weight on many different groups, so that no element of the resulting smoothed group-level distribution contains identifying information about any particular worker or establishment, eliminating disclosure risk.

### A7 Model Validation

The results presented in the preceding section consider relatively large, locally focused labor demand shocks, but the estimated surplus parameters \( \hat{\Theta}^{D-in-D} \) that underlie the simulations are identified from millions of quotidian worker job transitions stemming from natural job turnover and preference or skill changes over the life cycle that generates huge amounts of offsetting churn in the American labor market. Thus, one might reasonably wonder whether the parameters that govern ordinary worker flows remain valid when considering the response to sizable, locally targeted positive or negative shocks. To address this concern, in this section I describe and present results from a model validation exercise, in which parameters from a two-sided assignment model estimated on pre-shock ordinary worker flows are used to forecast the reallocation of workers that followed actual local economic shocks observed in the LEHD sample.

Specifically, 514 shocks to employment in a census tract experienced by a were identified in the LEHD sample that satisfied the following criteria: 1) the shock occurred in a sample state during the years 1996 - 2010; 2) at least 100 more or 100 fewer positions (and at most 3000) were filled in the chosen census tract than the year before; 3) the change in the number of positions constituted at least 10% and at most 100% of the total number of filled positions in the chosen census tract in the prior year; 4) The chosen tract featured at least 200 positions in the year prior to the shock; 5) no other tract in the same PUMA experienced an offsetting shock more than 50% as large as the shock to the chosen tract; 6) less than 50% of the change in number of positions filled in the year of the shock were offset by a shock to the same tract in the opposite direction the following year.

These criteria were chosen to ensure that a sufficient number of states would be reporting data in both the shock year and the prior year to properly capture the worker reallocation that ensured, that the shock was big enough to represent a meaningful stimulus or disruption to both the chosen tract and the surrounding area, and that the shock was sufficiently persistent that the possibility of a spurious reporting error by a large firm in the unemployment insurance data was unlikely to cause the “shock”.

To create a forecast of the worker reallocations that a given shock occurring in year \( y \) would engender, the full set of model parameters was estimated based on the nationwide sample of worker transitions between years \( y - 2 \) and \( y - 1 \), using the same procedures for
smoothing and aggregating types featuring distant locations described in Section 5.2. A counterfactual allocation was then generated by holding fixed the estimated surplus parameters but imposing the marginal distributions of origin and destination types from the pair of years capturing the shock, \( f^{y-1}(o) \) and \( h^y(d) \). Since the exact composition of the shock (as reflected in \( h^y(d) \)) is built into the forecast, the test of the model is the degree to which the particular flows of workers of different origin types to particular destination position types that resulted from the shock can be predicted.

The criterion I use to assess the accuracy of the forecast is the index of dissimilarity, which measures the percentage of predicted worker transitions that would need to be reassigned to a different transition group in order to perfectly match the distribution of actual worker transitions across groups. It is computed by summing the absolute differences across all transition groups \( g \) in the share of all transitions assigned to \( g \) both in the forecast and in the actual data: \( \sum_g |P[\hat{g}] - P(g)| \).

To help understand the sources of improvements and shortfalls in model fit, I also compute the index of dissimilarity between the true allocation and four alternative forecasts. The first is a standard parametric conditional logit specification, in which the probability that a random position of type \( d \) is filled by a worker whose transition would be assigned to group \( g \) is given by \( P[\hat{g}] = \frac{e^{X_y^g \lambda}}{\sum_{g'} e^{X_y^{g'} \lambda}} \), where \( X_y^g \) includes a substantial set of regressors constructed for year \( y \) that capture the kinds of predictors of joint surplus that researchers often use, and \( \lambda \) is the corresponding vector of parameters estimated from the relationship between the previous year’s \( P(g|d) \) and \( X_g \). The regressors include full sets of dummies for the following categorical variables: origin-destination distance bins using tract pathlength within PUMA, PUMA pathlength within state, and State pathlength between states, initial earnings quartile \( \times \) supersector dummies, initial earnings \( \times \) firm size quartile dummies, initial earnings \( \times \) firm average pay quartile dummies. The regressors also include an indicator for whether the group \( g \) is associated with job movers or stayers \((1(z(g) = 1))\), and the origin type frequency \( f(o(g)) \) interacted with the geographic category of the destination type associated with \( g \) (tract, PUMA, or state), interactions between \( f(o(g)) \) and indicators for whether \( d(g) \) represents the “nonemployment” position type, and dummies for whether the origin and destination types associated with transition group \( g \) are in the same PUMA and in the same state.

The second alternative forecast simply imposes that the conditional choice probability that existed between \( y - 2 \) and \( y - 1 \) also holds during the year of the shock, so that \( P(g) = \hat{P}^{y-1}(g|d)h^y(d) \). The third alternative forecast mimics the second, except that the smoothing procedure described in Section A6 is applied to the year \( y - 2 \) data prior to constructing \( \hat{P}^{y-1}(g|d) \). Like much research on either worker job search or firm job filling, all of these three alternative forecasts ignore the two-sided nature of the problem, and thus do not impose that the proposed allocation satisfies the marginal distribution of worker
origin types, $f^{y-1}(o)$. The fourth alternative forecast consists of a random matching in which the probability that a transition belongs to group $g$ is simply the product of the marginals of the associated origin and destination types, multiplied by the probability that a chosen worker would be a stayer vs. new hire: $P_g = f^{y-1}(o)S_{g|o,d}$. This forecast assumes each worker/position match is equally likely to take place, regardless of worker and establishment characteristics and locations. Note, though, that it does impose that the both the forecasted origin and destination marginal distributions match their actual counterparts.

Row 1 of Table A25 contains the results of this exercise. All entries consist of averages across the 514 shocks considered. The first five columns form the index of dissimilarity over all groups $g$ in the 19 state sample, while the second five columns only consider the allocation among groups $g$ featuring origin worker types associated with the same PUMA as the tract receiving the shock, so as to hone in on the local area most disrupted by the shock. The two-sided matching model, with parameters estimated from the previous period, would only need 7.1% of all worker transitions in the country to be reallocated to different transition groups to perfectly match the data, although 36.2% of the workers entering the period in the relevant PUMA were misallocated. Note though, that predicting the exact origin census tract and initial earnings quartile/nonemployment status of each worker hired separately for positions defined by tract/size/avg. pay/industry combinations is quite a tall order. Comparing across columns, we see that the parametric logit, despite over 100 regressors, performs considerably worse: nearly 43% of all U.S. transitions and 50.4% of transitions starting in the relevant PUMA would need to be reallocated to a different group to match the actual post-shock allocation that took place. Holding fixed the full prior year conditional choice probability distribution (cols. 3 and 8) performs slightly worse than the two-sided estimator within the target PUMA (37.5% misallocated), while adding the smoothing procedure improves the fit to 35.6%.

For many purposes, however, forecasting exactly the right origin and destination tracts of transitions may be less important than correctly assessing the degree to which the disruption dissipates farther from the shock. To this end, row 2 reports results in which groups featuring the same worker and establishment characteristics and origin and destination locations that belong to the same distance bin as each other are combined (using 42 bins), so that the index of dissimilarity is computed over a somewhat coarser set of possible transition groups. Only 21.1% of transitions are now misallocated by the two-sided forecast, with the two CCP forecasts following suit, suggesting that a substantial share of “incorrect” predictions might nonetheless be sufficiently accurate for most purposes. Furthermore, row 3 shows that combining groups featuring the same distance bins and worker earnings category but different establishment characteristics (size, avg. pay, and industry categories) reduces the index of dissimilarity to 5.3% for workers originating in the targeted PUMA, and below 1% nationally. Furthermore, the two-sided model outperforms the simpler smoothed and
unsmoothed CCP models at this level of aggregation (5.3% vs. 6.6% and 7.8%, respectively, within PUMA). This suggests that the two-sided matching model is doing a better job of matching the locations of job movers and stayers, but is slightly less effective at matching small differences in the destination establishment characteristics of the jobs to which workers move. Aggregating from 42 to 17 larger distance bins provides a slight improvement in accuracy, showing again that many “incorrect” predictions are nonetheless fairly accurate.

For other purposes, the primary goal of a forecast might be to properly predict the geographic and skill incidence of unemployment. To this end, row 5 computes the index of dissimilarity exclusively over the set of groups featuring the nonemployment destination type, so that the exercise is to predict the share of nonemployed workers that will originate from each combination of location and initial earnings/unemployment status. Using the full set of locations, the origin types of only 5.9% of workers who end up nonemployed would need to be altered in order for the two-sided prediction to match the allocation that actually occurred. Focusing on only the workers originally working (or most recently working) within the target PUMA increases this value to 14.6%. The two-sided estimator matches the performance of the CCP estimators within PUMA and outperforms them nationally. Aggregating locations into coarse distance bins shows that the two-sided predictions only badly predicts origin distance from the shock for 7.8% of workers originating in the PUMA (and 4.3% nationally) who end up unemployed, suggesting that it predicts quite well the geographic and skill incidence of unemployment following the shocks considered.

Finally, because many unemployed workers are long-term unemployed whose continued unemployment is easy to predict, the last two rows put the model to a more stringent challenge. Here, only employment-to-unemployment and unemployment-to-employment transitions are considered, so that the goal is to predict the scale and geographic and skill composition of transitions into and out of employment following a shock, ignoring workers who remain employed or remain unemployed in both years. When locations are aggregated to coarse distance bins (which still leaves 11*6 = 66 total transition groups), only 12.1% of E-to-UE and UE-to-E transitions among workers originating within the targeted PUMA are incorrectly predicted by the model, and only 4.4% nationally. Again, the two-sided model drastically outperforms the parametric model, outperforms the smoothed CCP model at both local and national levels, and outperforms the raw CCP model at the national level while nearly matching it at the local level.

Taken together, the model does quite a good job of predicting the reallocation of workers across job types and particularly across employment/unemployment status that follows major local labor market shocks.
### Table A1: Specifications for Alternative Counterfactual Labor Demand Shocks

<table>
<thead>
<tr>
<th>Spec. No.</th>
<th>Number of Jobs (or % of Tract’s Jobs)</th>
<th>Firm Avg. Earn. Quartile</th>
<th>Firm Size Quartile</th>
<th>Industry Supersector</th>
<th>Shock Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Information</td>
<td>Stimulus</td>
</tr>
<tr>
<td>2</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Information</td>
<td>Stimulus</td>
</tr>
<tr>
<td>3</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Information</td>
<td>Stimulus</td>
</tr>
<tr>
<td>4</td>
<td>500</td>
<td>4</td>
<td>4</td>
<td>Information</td>
<td>Stimulus</td>
</tr>
<tr>
<td>5</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Manufacturing</td>
<td>Stimulus</td>
</tr>
<tr>
<td>6</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Manufacturing</td>
<td>Stimulus</td>
</tr>
<tr>
<td>7</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Manufacturing</td>
<td>Stimulus</td>
</tr>
<tr>
<td>8</td>
<td>500</td>
<td>4</td>
<td>4</td>
<td>Manufacturing</td>
<td>Stimulus</td>
</tr>
<tr>
<td>9</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Trade/Trans./Utilities</td>
<td>Stimulus</td>
</tr>
<tr>
<td>10</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Trade/Trans./Utilities</td>
<td>Stimulus</td>
</tr>
<tr>
<td>11</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Trade/Trans./Utilities</td>
<td>Stimulus</td>
</tr>
<tr>
<td>12</td>
<td>500</td>
<td>4</td>
<td>4</td>
<td>Trade/Trans./Utilities</td>
<td>Stimulus</td>
</tr>
<tr>
<td>13</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Other Services</td>
<td>Stimulus</td>
</tr>
<tr>
<td>14</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Other Services</td>
<td>Stimulus</td>
</tr>
<tr>
<td>15</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Other Services</td>
<td>Stimulus</td>
</tr>
<tr>
<td>16</td>
<td>500</td>
<td>4</td>
<td>4</td>
<td>Other Services</td>
<td>Stimulus</td>
</tr>
<tr>
<td>17</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Education &amp; Health</td>
<td>Stimulus</td>
</tr>
<tr>
<td>18</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Education &amp; Health</td>
<td>Stimulus</td>
</tr>
<tr>
<td>19</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Education &amp; Health</td>
<td>Stimulus</td>
</tr>
<tr>
<td>20</td>
<td>500</td>
<td>4</td>
<td>4</td>
<td>Education &amp; Health</td>
<td>Stimulus</td>
</tr>
<tr>
<td>21</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Leisure &amp; Hospitality</td>
<td>Stimulus</td>
</tr>
<tr>
<td>22</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Leisure &amp; Hospitality</td>
<td>Stimulus</td>
</tr>
<tr>
<td>23</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Leisure &amp; Hospitality</td>
<td>Stimulus</td>
</tr>
<tr>
<td>24</td>
<td>500</td>
<td>4</td>
<td>4</td>
<td>Leisure &amp; Hospitality</td>
<td>Stimulus</td>
</tr>
<tr>
<td>25</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Government</td>
<td>Stimulus</td>
</tr>
<tr>
<td>26</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Government</td>
<td>Stimulus</td>
</tr>
<tr>
<td>27</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Government</td>
<td>Stimulus</td>
</tr>
<tr>
<td>28</td>
<td>500</td>
<td>4</td>
<td>4</td>
<td>Government</td>
<td>Stimulus</td>
</tr>
<tr>
<td>29</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Construction</td>
<td>Stimulus</td>
</tr>
<tr>
<td>30</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Construction</td>
<td>Stimulus</td>
</tr>
<tr>
<td>31</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Construction</td>
<td>Stimulus</td>
</tr>
<tr>
<td>32</td>
<td>500</td>
<td>4</td>
<td>4</td>
<td>Construction</td>
<td>Stimulus</td>
</tr>
<tr>
<td>33</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Information</td>
<td>Relocation</td>
</tr>
<tr>
<td>34</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Manufacturing</td>
<td>Relocation</td>
</tr>
<tr>
<td>35</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Trade/Trans./Utilities</td>
<td>Relocation</td>
</tr>
<tr>
<td>36</td>
<td>500</td>
<td>4</td>
<td>1</td>
<td>Information</td>
<td>Restr. Stim.</td>
</tr>
<tr>
<td>37</td>
<td>500</td>
<td>2</td>
<td>4</td>
<td>Manufacturing</td>
<td>Restr. Stim.</td>
</tr>
<tr>
<td>38</td>
<td>500</td>
<td>2</td>
<td>1</td>
<td>Trade/Trans./Utilities</td>
<td>Restr. Stim.</td>
</tr>
<tr>
<td>39</td>
<td>25%</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>Disaster</td>
</tr>
<tr>
<td>40</td>
<td>50%</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>Disaster</td>
</tr>
<tr>
<td>41</td>
<td>100%</td>
<td>All</td>
<td>All</td>
<td>All</td>
<td>Disaster</td>
</tr>
</tbody>
</table>
Table A2: Distribution of Destination Distances among Job Movers and Shares of All Workers by Distance Bin

<table>
<thead>
<tr>
<th>Distance Category</th>
<th>Share of Dest. Among Movers</th>
<th>Share of Initial Positions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same Tract</td>
<td>0.035</td>
<td>4.5E-05</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>0.076</td>
<td>2.6E-04</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>0.072</td>
<td>5.9E-04</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>0.138</td>
<td>0.003</td>
</tr>
<tr>
<td>1 PUMA</td>
<td>0.088</td>
<td>0.004</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>0.147</td>
<td>0.011</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>0.365</td>
<td>0.162</td>
</tr>
<tr>
<td>1 State Away</td>
<td>0.031</td>
<td>0.135</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>0.049</td>
<td>0.684</td>
</tr>
</tbody>
</table>

Notes: In column 1, the distance bins given by the row labels represent distances between origin and destination positions for workers who changed dominant jobs between 2010 and 2011. The cells in column 1 capture the share of all transitions whose origin-destination distance fell into the chosen bin. In column 2, these distance bins capture the distance between a worker’s dominant job and the census tract receiving a (simulated) labor demand shock. The cells capture the share of workers for whom the distance between their origin position and the targeted census tract fell into the chosen bin (averaged over 500 simulations featuring different target census tracts). “1 Tct Away”/“2 Tcts Away”/“3+ Tcts w/in PUMA” represent adjacent tracts, 2-tract pathlengths, or 3+ tract pathlengths within the same Public Use Microdata Area (PUMA), respectively. Likewise, “1 PUMA”, “2 PUMAs Away”, and “3+ PUMAs w/in State” represent adjacent PUMAs, 2 PUMA pathlengths, or 3+ PUMA pathlengths within the same state, respectively, and “1 State Away” and “2+ States Away” represent adjacent states and pathlengths of two or more states, respectively.
Table A3: Summary Statistics Describing Heterogeneity in the Spatial Scope of Labor Markets by Origin Worker Characteristics and Destination Establishment Characteristics

### Panel A: By Origin Worker Unemployment or Earnings Category

<table>
<thead>
<tr>
<th>Worker Subpop.</th>
<th># of Obs.</th>
<th>Unemp. to Unemp.</th>
<th>Emp. to Unemp.</th>
<th>Stay at Same Job</th>
<th>Move to New Job</th>
<th>Share of Transitions</th>
<th>Share of Transitions to New Jobs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Same PUMA</td>
<td>New PUMA, Same State</td>
</tr>
<tr>
<td>All Workers</td>
<td>23485000</td>
<td>0.042</td>
<td>0.043</td>
<td>0.703</td>
<td>0.213</td>
<td>0.320</td>
<td>0.600</td>
</tr>
<tr>
<td>Young (&lt;25) Unemp.</td>
<td>1018000</td>
<td>0.154</td>
<td>0.000</td>
<td>0.000</td>
<td>0.846</td>
<td>0.267</td>
<td>0.636</td>
</tr>
<tr>
<td>Older (≥ 25) Unemp.</td>
<td>1747000</td>
<td>0.469</td>
<td>0.000</td>
<td>0.000</td>
<td>0.531</td>
<td>0.287</td>
<td>0.600</td>
</tr>
<tr>
<td>1st Earn. Quart.</td>
<td>4947000</td>
<td>0.000</td>
<td>0.097</td>
<td>0.667</td>
<td>0.236</td>
<td>0.343</td>
<td>0.604</td>
</tr>
<tr>
<td>2nd Earn. Quart.</td>
<td>5028000</td>
<td>0.000</td>
<td>0.052</td>
<td>0.784</td>
<td>0.164</td>
<td>0.355</td>
<td>0.581</td>
</tr>
<tr>
<td>3rd Earn. Quart.</td>
<td>5121000</td>
<td>0.000</td>
<td>0.030</td>
<td>0.849</td>
<td>0.121</td>
<td>0.350</td>
<td>0.578</td>
</tr>
<tr>
<td>4th Earn. Quart.</td>
<td>5647000</td>
<td>0.000</td>
<td>0.020</td>
<td>0.873</td>
<td>0.107</td>
<td>0.331</td>
<td>0.583</td>
</tr>
</tbody>
</table>

### Panel B: By Destination Establishment Pay Quartile and Size Quartile

<table>
<thead>
<tr>
<th>Estab. Subpop.</th>
<th># of Obs.</th>
<th>Unemp. to Unemp.</th>
<th>Emp. to Unemp.</th>
<th>Stay at Same Job</th>
<th>Move to New Job</th>
<th>Share of Transitions</th>
<th>Share of Transitions to New Jobs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Same PUMA</td>
<td>New PUMA, Same State</td>
</tr>
<tr>
<td>1st Q. Avg. Earn.</td>
<td>5932000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.668</td>
<td>0.332</td>
<td>0.323</td>
<td>0.603</td>
</tr>
<tr>
<td>2nd Q. Avg. Earn.</td>
<td>5362000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.783</td>
<td>0.217</td>
<td>0.334</td>
<td>0.590</td>
</tr>
<tr>
<td>3rd Q. Avg. Earn.</td>
<td>4835000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.815</td>
<td>0.185</td>
<td>0.331</td>
<td>0.591</td>
</tr>
<tr>
<td>4th Q. Avg. Earn.</td>
<td>5382000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.819</td>
<td>0.181</td>
<td>0.292</td>
<td>0.612</td>
</tr>
<tr>
<td>1st Q. Size</td>
<td>2374000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.682</td>
<td>0.318</td>
<td>0.284</td>
<td>0.446</td>
</tr>
<tr>
<td>2nd Q. Size</td>
<td>2063000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.722</td>
<td>0.278</td>
<td>0.344</td>
<td>0.592</td>
</tr>
<tr>
<td>3rd Q. Size</td>
<td>2091000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.727</td>
<td>0.273</td>
<td>0.326</td>
<td>0.625</td>
</tr>
<tr>
<td>4th Q. Size</td>
<td>15003000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.793</td>
<td>0.207</td>
<td>0.324</td>
<td>0.634</td>
</tr>
</tbody>
</table>

### Panel C: By Destination Establishment Industry

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Same PUMA</td>
<td>New PUMA, Same State</td>
</tr>
<tr>
<td>Nat. Resources</td>
<td>3699000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.692</td>
<td>0.308</td>
<td>0.496</td>
<td>0.406</td>
</tr>
<tr>
<td>Construction</td>
<td>9747000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.741</td>
<td>0.259</td>
<td>0.319</td>
<td>0.586</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>19292000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.840</td>
<td>0.160</td>
<td>0.398</td>
<td>0.530</td>
</tr>
<tr>
<td>Whole/Retail/Trans.</td>
<td>4382000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.764</td>
<td>0.236</td>
<td>0.280</td>
<td>0.635</td>
</tr>
<tr>
<td>Information</td>
<td>5391000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.799</td>
<td>0.201</td>
<td>0.318</td>
<td>0.599</td>
</tr>
<tr>
<td>Financial Activities</td>
<td>1287800</td>
<td>0.000</td>
<td>0.000</td>
<td>0.769</td>
<td>0.231</td>
<td>0.285</td>
<td>0.648</td>
</tr>
<tr>
<td>Prof. Bus. Services</td>
<td>3020000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.692</td>
<td>0.308</td>
<td>0.275</td>
<td>0.642</td>
</tr>
<tr>
<td>Ed. Health</td>
<td>4859000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.815</td>
<td>0.185</td>
<td>0.363</td>
<td>0.566</td>
</tr>
<tr>
<td>Leis. &amp; Hosp.</td>
<td>2269000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.670</td>
<td>0.330</td>
<td>0.328</td>
<td>0.583</td>
</tr>
<tr>
<td>Oth. Serv.</td>
<td>8413000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.744</td>
<td>0.256</td>
<td>0.330</td>
<td>0.604</td>
</tr>
<tr>
<td>Government</td>
<td>10599000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.910</td>
<td>0.090</td>
<td>0.393</td>
<td>0.550</td>
</tr>
</tbody>
</table>

### Table A4: Assessing the Impact of Stimulus Packages at Different Distances from Focal Tract Across Several Outcomes

Stimuli Consist of 500 New Jobs (Averages Across All Stimulus Compositions)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Tract</td>
<td>0.015 (1.1E-04)</td>
<td>0.031 (9.2E-05)</td>
<td>0.003 (1.7E-05)</td>
<td>0.005 (1.4E-05)</td>
<td>1045 (25)</td>
<td>0.005 (1.9E-05)</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>0.005 (3.3E-05)</td>
<td>0.057 (1.5E-04)</td>
<td>9.4E-04 (4.7E-06)</td>
<td>0.012 (2.6E-05)</td>
<td>395 (7)</td>
<td>0.010 (3.5E-05)</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>0.002 (1.1E-05)</td>
<td>0.067 (1.4E-04)</td>
<td>6.0E-04 (2.2E-06)</td>
<td>0.017 (3.2E-05)</td>
<td>278 (3)</td>
<td>0.014 (3.9E-05)</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>0.001 (6.7E-06)</td>
<td>0.138 (1.8E-04)</td>
<td>3.8E-04 (1.3E-06)</td>
<td>0.047 (6.1E-05)</td>
<td>164 (1)</td>
<td>0.041 (7.9E-05)</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>5.5E-04 (1.2E-06)</td>
<td>0.096 (1.6E-04)</td>
<td>2.6E-04 (4.4E-07)</td>
<td>0.046 (6.9E-05)</td>
<td>164 (0.7)</td>
<td>0.042 (8.5E-05)</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>2.8E-04 (4.5E-07)</td>
<td>0.149 (1.8E-04)</td>
<td>1.8E-04 (2.7E-07)</td>
<td>0.095 (1.1E-04)</td>
<td>143 (0.4)</td>
<td>0.083 (1.3E-04)</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>5.3E-05 (2.2E-07)</td>
<td>0.355 (4.1E-04)</td>
<td>6.1E-05 (1.7E-07)</td>
<td>0.433 (4.7E-04)</td>
<td>109 (0.5)</td>
<td>0.390 (5.0E-04)</td>
</tr>
<tr>
<td>1 State Away</td>
<td>8.3E-06 (2.0E-08)</td>
<td>0.053 (1.3E-04)</td>
<td>2.0E-05 (2.0E-08)</td>
<td>0.128 (2.0E-04)</td>
<td>89 (0.0)</td>
<td>0.132 (2.1E-04)</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>1.6E-06 (2.4E-09)</td>
<td>0.053 (7.9E-05)</td>
<td>6.7E-06 (4.1E-04)</td>
<td>0.217 (1.8E-04)</td>
<td>85 (0.0)</td>
<td>0.284 (4.6E-04)</td>
</tr>
</tbody>
</table>

Notes: The column labeled “Prob. of Stim. Job” indicates the probability that a randomly chosen worker in the row subgroup will receive one of the 500 new positions generated by the simulated stimulus package. The column labeled “Change in P(Employed)” indicates the change in the probability that a randomly chosen worker in the row subgroup will be employed in the destination year as a consequence of the simulated stimulus package. The column labeled “Avg. Welfare Change” indicates the change in job-related welfare (scaled to be equivalent to $ of 2010 annual earnings) that a randomly chosen worker in the subgroup indicated by the row label will experience as a consequence of the simulated stimulus package. The columns labeled “Share of Stim. Jobs”, “Share of Emp. Gains” and “Share of Wel. Gains” indicate the share of all stimulus jobs and total employment and welfare gains, respectively, generated by the simulated stimulus package that accrue to workers in the subgroup indicated by the row label.

“Target Tract” indicates that the worker’s origin establishment was in the tract receiving the stimulus package. “1/2/3+ Tct(s) Away” indicates that the origin establishment was one, two, or 3 or more tracts away (by tract pathlength) within the same PUMA. “1/2+ PUMAs Away” and “1/2+ States Away” indicate the PUMA pathlength (if within the same state) and state pathlength (if in different states), respectively.

Standard errors are provided in parentheses, and are based on the sampling distribution among the sample of 500 target tracts simulated for each stimulus package specification.
Table A5: Average of Each Incidence Measure by Distance from Target Tract Across All Stimulus Packages, Measured in Miles - Each Column Averages Results 500 Simulations Featuring 500 Different Target Census Tracts)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Within 1 Mile</td>
<td>480</td>
<td>0.009</td>
<td>8.0E-04</td>
<td>0.008</td>
<td>0.004</td>
<td>0.036</td>
</tr>
<tr>
<td>1-2 Miles Away</td>
<td>230</td>
<td>0.014</td>
<td>4.4E-04</td>
<td>0.011</td>
<td>0.002</td>
<td>0.046</td>
</tr>
<tr>
<td>3-5 Miles Away</td>
<td>210</td>
<td>0.041</td>
<td>2.9E-04</td>
<td>0.037</td>
<td>9.3E-04</td>
<td>0.117</td>
</tr>
<tr>
<td>6-11 Miles Away</td>
<td>226</td>
<td>0.052</td>
<td>1.9E-04</td>
<td>0.047</td>
<td>5.1E-04</td>
<td>0.124</td>
</tr>
<tr>
<td>11-26 Miles Away</td>
<td>266</td>
<td>0.087</td>
<td>1.7E-04</td>
<td>0.077</td>
<td>3.7E-04</td>
<td>0.167</td>
</tr>
<tr>
<td>26-50 Miles Away</td>
<td>152</td>
<td>0.044</td>
<td>1.6E-04</td>
<td>0.041</td>
<td>2.8E-04</td>
<td>0.072</td>
</tr>
<tr>
<td>51-100 Miles Away</td>
<td>107</td>
<td>0.055</td>
<td>1.6E-04</td>
<td>0.055</td>
<td>1.7E-04</td>
<td>0.061</td>
</tr>
<tr>
<td>101-250 Miles Away</td>
<td>82</td>
<td>0.165</td>
<td>5.4E-05</td>
<td>0.159</td>
<td>3.5E-05</td>
<td>0.103</td>
</tr>
<tr>
<td>&gt;250 Miles Away</td>
<td>66</td>
<td>0.533</td>
<td>1.3E-05</td>
<td>0.565</td>
<td>6.4E-06</td>
<td>0.275</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the outcomes in the column labels. The row labels define subpopulations of workers for whom the distance between the establishment associated with their origin dominant jobs and the census tract receiving the simulated stimulus package fell in the listed distance bin.
Table A6: Change in Probability of Employment due to Stimulus for a Randomly Chosen Individual at Different Distances from Focal Tract: Stimuli Consist of 500 New Jobs at Firms in Alternative Industries (Averaged Across Firm Size/Firm Average Earnings Combinations)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Tract</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
<td>0.002</td>
<td>0.003</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(1.7E-05)</td>
<td>(1.8E-05)</td>
<td>(1.9E-05)</td>
<td>(2.1E-05)</td>
<td>(2.2E-05)</td>
<td>(1.8E-05)</td>
<td>(1.9E-05)</td>
<td>(1.8E-05)</td>
<td>(1.9E-05)</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>9.4E-04</td>
<td>9.4E-04</td>
<td>9.8E-04</td>
<td>8.6E-04</td>
<td>0.001</td>
<td>9.5E-04</td>
<td>9.6E-04</td>
<td>9.1E-04</td>
<td>9.2E-04</td>
</tr>
<tr>
<td></td>
<td>(4.7E-06)</td>
<td>(5.0E-06)</td>
<td>(5.1E-06)</td>
<td>(5.2E-06)</td>
<td>(6.2E-06)</td>
<td>(5.0E-06)</td>
<td>(5.3E-06)</td>
<td>(4.5E-06)</td>
<td>(6.5E-06)</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>6.0E-04</td>
<td>5.8E-04</td>
<td>6.2E-04</td>
<td>5.4E-04</td>
<td>6.7E-04</td>
<td>6.4E-04</td>
<td>6.0E-04</td>
<td>5.9E-04</td>
<td>6.0E-04</td>
</tr>
<tr>
<td></td>
<td>(2.2E-06)</td>
<td>(2.2E-06)</td>
<td>(2.5E-06)</td>
<td>(2.6E-06)</td>
<td>(2.7E-06)</td>
<td>(2.5E-06)</td>
<td>(2.7E-06)</td>
<td>(2.3E-06)</td>
<td>(2.6E-06)</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>3.8E-04</td>
<td>3.6E-04</td>
<td>3.6E-04</td>
<td>3.5E-04</td>
<td>3.9E-04</td>
<td>4.3E-04</td>
<td>3.9E-04</td>
<td>3.7E-04</td>
<td>3.7E-04</td>
</tr>
<tr>
<td></td>
<td>(1.3E-06)</td>
<td>(1.3E-06)</td>
<td>(1.3E-06)</td>
<td>(1.7E-06)</td>
<td>(1.4E-06)</td>
<td>(1.6E-06)</td>
<td>(1.8E-06)</td>
<td>(1.3E-06)</td>
<td>(1.3E-06)</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>2.6E-04</td>
<td>2.5E-04</td>
<td>2.6E-04</td>
<td>2.5E-04</td>
<td>2.8E-04</td>
<td>2.7E-04</td>
<td>2.6E-04</td>
<td>2.6E-04</td>
<td>2.6E-04</td>
</tr>
<tr>
<td></td>
<td>(4.4E-07)</td>
<td>(4.3E-07)</td>
<td>(4.8E-07)</td>
<td>(4.8E-07)</td>
<td>(5.2E-07)</td>
<td>(6.3E-07)</td>
<td>(5.0E-07)</td>
<td>(4.6E-07)</td>
<td>(6.3E-07)</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>1.8E-04</td>
<td>1.8E-04</td>
<td>1.8E-04</td>
<td>1.8E-04</td>
<td>1.8E-04</td>
<td>1.7E-04</td>
<td>1.8E-04</td>
<td>1.8E-04</td>
<td>1.8E-04</td>
</tr>
<tr>
<td></td>
<td>(2.7E-07)</td>
<td>(2.7E-07)</td>
<td>(2.9E-07)</td>
<td>(3.0E-07)</td>
<td>(3.1E-07)</td>
<td>(2.8E-07)</td>
<td>(3.0E-07)</td>
<td>(3.0E-07)</td>
<td>(3.2E-07)</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>6.1E-05</td>
<td>6.2E-05</td>
<td>6.1E-05</td>
<td>6.2E-05</td>
<td>6.1E-05</td>
<td>6.0E-05</td>
<td>6.0E-05</td>
<td>6.2E-05</td>
<td>6.1E-05</td>
</tr>
<tr>
<td></td>
<td>(1.7E-07)</td>
<td>(1.8E-07)</td>
<td>(1.8E-07)</td>
<td>(1.7E-07)</td>
<td>(2.0E-07)</td>
<td>(1.7E-07)</td>
<td>(1.7E-07)</td>
<td>(1.9E-07)</td>
<td>(1.7E-07)</td>
</tr>
<tr>
<td>1 State Away</td>
<td>2.0E-05</td>
<td>2.0E-05</td>
<td>2.1E-05</td>
<td>2.0E-05</td>
<td>1.9E-05</td>
<td>2.0E-05</td>
<td>2.0E-05</td>
<td>2.0E-05</td>
<td>2.1E-05</td>
</tr>
<tr>
<td></td>
<td>(2.0E-08)</td>
<td>(2.0E-08)</td>
<td>(2.4E-08)</td>
<td>(2.1E-08)</td>
<td>(2.1E-08)</td>
<td>(2.6E-08)</td>
<td>(2.2E-08)</td>
<td>(2.2E-08)</td>
<td>(2.5E-08)</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>6.7E-06</td>
<td>6.7E-06</td>
<td>6.8E-06</td>
<td>6.8E-06</td>
<td>6.6E-06</td>
<td>6.7E-06</td>
<td>7.0E-06</td>
<td>6.6E-06</td>
<td>6.7E-06</td>
</tr>
<tr>
<td></td>
<td>(4.1E-09)</td>
<td>(4.7E-09)</td>
<td>(5.0E-09)</td>
<td>(5.1E-09)</td>
<td>(4.9E-09)</td>
<td>(5.0E-09)</td>
<td>(5.1E-09)</td>
<td>(4.2E-09)</td>
<td>(5.4E-09)</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the row labels. Each entry provides the average change in the probability of being employed in the destination year attributable to a 500 job stimulus package for workers whose distance between their origin jobs and the census tract receiving the stimulus package placed them in the distance bin indicated in the row label. Different columns consider average employment impacts from stimuli featuring jobs with establishments in the same industry supersector but in different quartiles of the establishment-level employment and average worker earnings distributions. Standard errors are provided in parentheses, and are based on the sampling distribution among the sample of 500 target tracts simulated for each stimulus package specification. “Avg.”: Average employment change across all 32 stimulus packages considered (and all 500 target tracts for each stimulus package specification. “Info”: Information. “Manu.”: Manufacturing. “Trd./Tns.”: Trade/Transportation/Utilities. “Oth. Serv.”: Other Services (includes repair, laundry, security, personal services). “Ed./Hlth”: Education and Healthcare. “Lei/Hosp.”: Leisure and Hospitality. “Gov.”: Government. “Const.”: Construction.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Tract</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.007</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(1.4E-05)</td>
<td>(1.4E-05)</td>
<td>(1.6E-05)</td>
<td>(2.0E-05)</td>
<td>(1.5E-05)</td>
<td>(1.5E-05)</td>
<td>(1.5E-05)</td>
<td>(1.8E-05)</td>
<td>(1.8E-05)</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>0.012</td>
<td>0.012</td>
<td>0.011</td>
<td>0.012</td>
<td>0.012</td>
<td>0.012</td>
<td>0.011</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(2.6E-05)</td>
<td>(2.8E-05)</td>
<td>(3.0E-05)</td>
<td>(3.4E-05)</td>
<td>(3.0E-05)</td>
<td>(3.0E-05)</td>
<td>(3.0E-05)</td>
<td>(2.6E-05)</td>
<td>(2.7E-05)</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>0.017</td>
<td>0.016</td>
<td>0.017</td>
<td>0.015</td>
<td>0.018</td>
<td>0.018</td>
<td>0.017</td>
<td>0.016</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(3.2E-05)</td>
<td>(3.3E-05)</td>
<td>(3.7E-05)</td>
<td>(3.2E-05)</td>
<td>(4.0E-05)</td>
<td>(4.0E-05)</td>
<td>(3.8E-05)</td>
<td>(3.4E-05)</td>
<td>(3.7E-05)</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>0.047</td>
<td>0.047</td>
<td>0.047</td>
<td>0.043</td>
<td>0.051</td>
<td>0.050</td>
<td>0.048</td>
<td>0.047</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>(6.1E-05)</td>
<td>(6.5E-05)</td>
<td>(6.7E-05)</td>
<td>(6.5E-05)</td>
<td>(7.1E-05)</td>
<td>(7.0E-05)</td>
<td>(6.7E-05)</td>
<td>(6.5E-05)</td>
<td>(7.1E-05)</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>0.046</td>
<td>0.044</td>
<td>0.045</td>
<td>0.044</td>
<td>0.050</td>
<td>0.048</td>
<td>0.045</td>
<td>0.045</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>(6.9E-05)</td>
<td>(7.2E-05)</td>
<td>(7.5E-05)</td>
<td>(7.9E-05)</td>
<td>(7.7E-05)</td>
<td>(7.7E-05)</td>
<td>(7.7E-05)</td>
<td>(7.1E-05)</td>
<td>(8.0E-05)</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>0.095</td>
<td>0.094</td>
<td>0.094</td>
<td>0.095</td>
<td>0.099</td>
<td>0.095</td>
<td>0.093</td>
<td>0.095</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>(1.1E-04)</td>
<td>(1.1E-04)</td>
<td>(1.1E-04)</td>
<td>(1.2E-04)</td>
<td>(1.2E-04)</td>
<td>(1.1E-04)</td>
<td>(1.1E-04)</td>
<td>(1.2E-04)</td>
<td>(1.2E-04)</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>0.433</td>
<td>0.438</td>
<td>0.431</td>
<td>0.439</td>
<td>0.429</td>
<td>0.427</td>
<td>0.428</td>
<td>0.440</td>
<td>0.430</td>
</tr>
<tr>
<td></td>
<td>(4.7E-04)</td>
<td>(4.9E-04)</td>
<td>(4.9E-04)</td>
<td>(4.9E-04)</td>
<td>(4.7E-04)</td>
<td>(4.8E-04)</td>
<td>(4.8E-04)</td>
<td>(4.8E-04)</td>
<td>(4.8E-04)</td>
</tr>
<tr>
<td>1 State Away</td>
<td>0.128</td>
<td>0.128</td>
<td>0.131</td>
<td>0.129</td>
<td>0.123</td>
<td>0.130</td>
<td>0.127</td>
<td>0.127</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>(2.0E-04)</td>
<td>(2.1E-04)</td>
<td>(2.1E-04)</td>
<td>(2.1E-04)</td>
<td>(2.0E-04)</td>
<td>(2.1E-04)</td>
<td>(2.1E-04)</td>
<td>(2.0E-04)</td>
<td>(2.2E-04)</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>0.217</td>
<td>0.216</td>
<td>0.218</td>
<td>0.220</td>
<td>0.212</td>
<td>0.216</td>
<td>0.225</td>
<td>0.213</td>
<td>0.216</td>
</tr>
<tr>
<td></td>
<td>(1.8E-04)</td>
<td>(2.0E-04)</td>
<td>(2.1E-04)</td>
<td>(2.0E-04)</td>
<td>(2.0E-04)</td>
<td>(2.1E-04)</td>
<td>(1.9E-04)</td>
<td>(2.2E-04)</td>
<td>(2.2E-04)</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the row labels. See Table A6 for expanded definitions of industry supersectors listed in column labels.
Table A8: Change in Probability of Employment and Share of Nationwide Employment Gains From New Stimulus Positions for a Randomly Chosen Individual at Different Distances from Focal Tract: Stimuli Consist of 500 New Positions in Alternative Combinations of Firm Size Quartile/Firm Average Pay Quartile (Averaged Across Industry Supersectors)

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>Change in P(Employed)</th>
<th>Share of Emp. Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sm./Low</td>
<td>Lg./Low</td>
</tr>
<tr>
<td>Target Tract</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>6.2E-04</td>
<td>6.9E-04</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>3.4E-04</td>
<td>3.9E-04</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>2.5E-04</td>
<td>2.8E-04</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>1.7E-04</td>
<td>1.9E-04</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>5.4E-05</td>
<td>5.9E-05</td>
</tr>
<tr>
<td>1 State Away</td>
<td>2.2E-05</td>
<td>1.8E-05</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>7.1E-06</td>
<td>6.1E-06</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the row labels. The first four columns capture the average change in the probability of being employed in the destination year attributable to a 500 job stimulus package for workers whose distance between their origin jobs and the census tract receiving the stimulus package place them in the distance bin indicated in the row label. The last four columns capture the share of all stimulus-driven employment gains accruing to workers in each distance bin. Different columns consider average employment impacts from stimuli featuring jobs with establishments from different combinations of firm size quartile and firm average worker earnings quartile in the respective nationwide establishment-level distributions. Each column averages results across 8 stimulus packages featuring jobs with establishments in the same firm size quartile/firm average pay quartile combination but in different industry suppersectors (as well as simulated 500 simulations for each stimulus package specification featuring different target census tracts). “Sm./Low”: The 500 stimulus jobs are generated by establishments whose employment levels place them in the smallest quartile of firms and whose average worker pay levels place them in the 2nd smallest quartile of firms. “Lg./Low”: The 500 stimulus jobs are generated by establishments whose employment levels place them in the largest quartile of firms and whose average worker pay levels place them in the 2nd smallest quartile of firms. “Sm./Hi”: The 500 stimulus jobs are generated by establishments whose employment levels place them in the smallest quartile of firms and whose average worker pay levels place them in the highest quartile of firms. “Lg./Hi”: The 500 stimulus jobs are generated by establishments whose employment levels place them in the largest quartile of firms and whose average worker pay levels place them in the highest quartile of firms.
Table A9: Expected Change in Utility and Share of Nationwide Utility Gains from New Stimulus Positions for a Randomly Chosen Individual at Different Distances from Focal Tract: Stimuli Consist of 500 New Positions in Alternative Combinations of Firm Size Quartile/Firm Average Pay Quartile (Averaged Across Industry Supersectors)

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>Avg. Welfare Change ($)</th>
<th>Share of Welfare Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sm./Low</td>
<td>Lg./Low</td>
</tr>
<tr>
<td>Target Tract</td>
<td>1197</td>
<td>1032</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>400</td>
<td>388</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>279</td>
<td>282</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>188</td>
<td>196</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>174</td>
<td>184</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>142</td>
<td>147</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>122</td>
<td>126</td>
</tr>
<tr>
<td>1 State Away</td>
<td>91</td>
<td>90</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>87</td>
<td>86</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the row labels. See Table A8 for expanded definitions of column labels. The first four columns capture the average change in job-related welfare (scaled to be equivalent to $ of 2010 annual earnings) in the destination year attributable to a 500 job stimulus package for workers whose distance between their origin jobs and the census tract receiving the stimulus package place them in the distance bin indicated in the row label. The last four columns capture the share of all stimulus-driven welfare gains accruing to workers in each distance bin. Each column averages results across 8 stimulus packages featuring jobs with establishments in the same firm size quartile/firm average pay quartile combination but in different industry supersectors (as well as simulated 500 simulations for each stimulus package specification featuring different target census tracts).
Table A10: Shares of Additional Employment and Utility Produced by Stimulus among Workers Initially Employed (or Unemployed) at Different Initial Earnings Quartiles (or Unemployment): Stimuli Consist of 500 New Jobs at Firms in Different Firm Size/Firm Average Earnings Quartiles (Averaged across Different Industries)

<table>
<thead>
<tr>
<th>Earnings Category</th>
<th>Share of Employment Gains</th>
<th>Share of Welfare Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avg.</td>
<td>Sm./Low</td>
</tr>
<tr>
<td>UE ≤ Age 25</td>
<td>0.089</td>
<td>0.087</td>
</tr>
<tr>
<td>UE &gt; Age 25</td>
<td>0.322</td>
<td>0.315</td>
</tr>
<tr>
<td>1st Quartile</td>
<td>0.261</td>
<td>0.269</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>0.154</td>
<td>0.159</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>0.096</td>
<td>0.096</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>0.076</td>
<td>0.071</td>
</tr>
</tbody>
</table>

Notes: See Table A8 for expanded definitions of column labels. The first four columns capture the average change in job-related welfare (scaled to be equivalent to $ of 2010 annual earnings) in the destination year attributable to a 500 job stimulus package for workers whose employment status or earnings in the origin year places them in the earnings/employment category listed by the row label. The last four columns capture the share of all stimulus-driven welfare gains accruing to workers in each earnings/employment category. Each column averages results across 8 stimulus packages featuring jobs with establishments in the same firm size quartile/firm average pay quartile combination but in different industry supersectors (as well as simulated 500 simulations for each stimulus package specification featuring different target census tracts). “UE ≤ Age 25”: Workers who were unemployed in the origin year (defined as no full quarter of work with >$2,000 in earnings at any establishment) and who were 25 years old or younger. “UE > Age 25”: Workers who were unemployed in the origin year and who were more than 25 years old. “1st/2nd/3rd/4th Quartile”: Workers whose average earnings among full quarters worked at their dominant job in the origin year placed them in the 1st/2nd/3rd/4th quartile of the 2010 annual earnings distribution for the sample states.
Table A11: Change in Probability of Employment due to Stimulus for a Randomly Chosen Individual at Different Combinations of Initial Earnings Quartile (or Nonemployment) and Distance from Focal Tract: Averaged Across All Stimulus Specifications Featuring 500 New Jobs

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>UE ≤ 25</th>
<th>UE &gt; 25</th>
<th>1st Q.</th>
<th>2nd Q.</th>
<th>3rd Q.</th>
<th>4th Q.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Tract</td>
<td>3.9E-03</td>
<td>1.6E-02</td>
<td>3.2E-03</td>
<td>1.8E-03</td>
<td>9.9E-04</td>
<td>6.3E-04</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>1.7E-03</td>
<td>4.3E-03</td>
<td>1.2E-03</td>
<td>7.2E-04</td>
<td>4.4E-04</td>
<td>2.7E-04</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>1.1E-03</td>
<td>2.7E-03</td>
<td>7.8E-04</td>
<td>4.5E-04</td>
<td>2.7E-04</td>
<td>1.8E-04</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>6.2E-04</td>
<td>1.5E-03</td>
<td>4.5E-04</td>
<td>2.7E-04</td>
<td>1.6E-04</td>
<td>9.6E-05</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>4.4E-04</td>
<td>1.1E-03</td>
<td>3.3E-04</td>
<td>2.0E-04</td>
<td>1.2E-04</td>
<td>7.7E-05</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>3.0E-04</td>
<td>6.8E-04</td>
<td>2.3E-04</td>
<td>1.4E-04</td>
<td>8.6E-05</td>
<td>5.5E-05</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>1.1E-04</td>
<td>2.1E-04</td>
<td>7.3E-05</td>
<td>4.5E-05</td>
<td>2.8E-05</td>
<td>1.6E-05</td>
</tr>
<tr>
<td>1 State Away</td>
<td>3.9E-05</td>
<td>8.1E-05</td>
<td>2.7E-05</td>
<td>1.5E-05</td>
<td>9.2E-06</td>
<td>6.3E-06</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>1.5E-05</td>
<td>3.0E-05</td>
<td>8.5E-06</td>
<td>4.7E-06</td>
<td>2.8E-06</td>
<td>2.0E-06</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the row labels. See Table A10 for expanded definitions of the column labels. Each cell contains the average change in the probability of employment in the destination year generated by a 500 job stimulus for workers whose distance between their origin establishment and the census tract receiving the stimulus package placed them in the distance bin indicated in the row label and whose employment status or earnings in the origin year placed them in the earnings/employment category listed by the column label. Each cell averages results across 32 stimulus packages featuring new jobs with establishments with different combinations of industry supersector, firm size quartile, and firm average pay quartile. Results are further averaged across 500 simulations for each of the 32 stimulus package specifications featuring different target census tracts.
Table A12: Expected Welfare Gain From New Stimulus Positions Among Workers Initially Employed at Different Combinations of Initial Earnings Quartile (or Nonemployed) and Distance from Focal Tract: Averaged Across All Stimulus Specifications Featuring 500 New Jobs

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>UE ≤ 25</th>
<th>UE &gt; 25</th>
<th>1st Q.</th>
<th>2nd Q.</th>
<th>3rd Q.</th>
<th>4th Q.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Tract</td>
<td>620</td>
<td>1165</td>
<td>999</td>
<td>1063</td>
<td>1104</td>
<td>1242</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>339</td>
<td>331</td>
<td>384</td>
<td>398</td>
<td>404</td>
<td>451</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>262</td>
<td>244</td>
<td>269</td>
<td>282</td>
<td>285</td>
<td>300</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>184</td>
<td>179</td>
<td>187</td>
<td>194</td>
<td>194</td>
<td>199</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>175</td>
<td>163</td>
<td>167</td>
<td>175</td>
<td>179</td>
<td>182</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>142</td>
<td>129</td>
<td>140</td>
<td>144</td>
<td>147</td>
<td>149</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>123</td>
<td>114</td>
<td>121</td>
<td>123</td>
<td>124</td>
<td>125</td>
</tr>
<tr>
<td>1 State Away</td>
<td>89</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>89</td>
<td>89</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>85</td>
<td>84</td>
<td>85</td>
<td>85</td>
<td>85</td>
<td>85</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the row labels. See Table A10 for expanded definitions of the column labels. Each cell contains the average job-related welfare gain (scaled to be equivalent to $ of 2010 annual earnings) generated by a 500 job stimulus for workers whose distance between their origin establishment and the census tract receiving the stimulus package placed them in the distance bin indicated in the row label and whose employment status or earnings in the origin year placed them in the earnings/employment category listed by the column label. Each cell averages results across 32 stimulus packages featuring new jobs with establishments with different combinations of industry supersector, firm size quartile, and firm average pay quartile. Results are further averaged across 500 simulations for each of the 32 stimulus package specifications featuring different target census tracts.
Table A13: Expected Job-Related Welfare Gain From New Stimulus Positions Among Workers Initially Employed in the Focal Tract at Different Earnings Quintiles (or Unemployed) by Industry Supersector (Averaged Across Firm Size/Firm Average Earnings Combinations)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>UE ≤ Age 25</td>
<td>620</td>
<td>515</td>
<td>655</td>
<td>632</td>
<td>757</td>
<td>561</td>
<td>823</td>
<td>551</td>
</tr>
<tr>
<td>UE &gt; Age 25</td>
<td>1165</td>
<td>952</td>
<td>1090</td>
<td>1097</td>
<td>1413</td>
<td>1102</td>
<td>1374</td>
<td>1000</td>
</tr>
<tr>
<td>1st Quartile</td>
<td>999</td>
<td>982</td>
<td>993</td>
<td>876</td>
<td>1282</td>
<td>968</td>
<td>1048</td>
<td>982</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>1063</td>
<td>1015</td>
<td>1097</td>
<td>958</td>
<td>1407</td>
<td>986</td>
<td>960</td>
<td>1110</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>1104</td>
<td>1062</td>
<td>1082</td>
<td>1017</td>
<td>1484</td>
<td>1001</td>
<td>969</td>
<td>1189</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>1242</td>
<td>1335</td>
<td>1283</td>
<td>1192</td>
<td>1640</td>
<td>1025</td>
<td>1127</td>
<td>1195</td>
</tr>
</tbody>
</table>

Notes: See Table A10 for expanded definitions of distance bins captured by the row labels. See Table A6 for expanded definitions of the industry supersectors captured by the column labels. Each cell contains the average job-related welfare gain (scaled to be equivalent to $ of 2010 annual earnings) generated by a 500 job stimulus for workers initially employed (or most recently employed) in the focal tract whose employment status or earnings in the origin year placed them in the earnings/employment category listed by the row label. Each column averages results across four stimulus packages featuring jobs with establishments in the same industry supersector but in different quartiles of the establishment-level employment and average worker earnings distributions. Results are further averaged across 500 simulations featuring different target census tracts for each of the stimulus package specifications.
Table A14: Expected Change in Utility From New Stimulus Positions Among Workers Initially Employed in the Focal Tract at Different Earnings Quintiles (or Nonemployed) by Firm Size Quartile/Firm Average Pay Quartile Combination (Averaged Across Industry Supersectors)

<table>
<thead>
<tr>
<th>Earnings Firm Size/Pay Level Combination</th>
<th>Sm./Low</th>
<th>Lg./Low</th>
<th>Sm./Hi</th>
<th>Lg./Hi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quintile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NE ≤ Age 25</td>
<td>625</td>
<td>802</td>
<td>413</td>
<td>640</td>
</tr>
<tr>
<td>NE &gt; Age 25</td>
<td>1169</td>
<td>1297</td>
<td>1015</td>
<td>1181</td>
</tr>
<tr>
<td>1st Quartile</td>
<td>1243</td>
<td>1130</td>
<td>825</td>
<td>800</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>1341</td>
<td>1102</td>
<td>934</td>
<td>874</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>1272</td>
<td>997</td>
<td>1134</td>
<td>1012</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>1081</td>
<td>872</td>
<td>1728</td>
<td>1287</td>
</tr>
</tbody>
</table>

Notes: See Table A10 for expanded definitions of employment status/earnings quartile categories captured by the row labels. See Table A8 for expanded definitions of the establishment size/avg. pay combinations captured by the column labels. Each cell contains the average job-related welfare gain (scaled to be equivalent to $ of 2010 annual earnings) generated by a 500 job stimulus for workers initially employed (or most recently employed) in the focal tract whose employment status or earnings in the origin year placed them in the earnings/employment category listed by the row label. Each column averages results from eight stimuli that feature jobs with establishments from different industry supersectors but the same quartiles of the establishment-level employment and average worker earnings distributions (indicated by the column label). Results are further averaged across 500 simulations featuring different target census tracts for each of the stimulus package specifications.
Table A15: Heterogeneity in Average Welfare Gain and Share of Total Welfare Gains by Distance from Focal Tract Across Focal Tracts of Varying Population and Employment Size

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>Avg. Welfare Gain ($)</th>
<th>Share of Welfare Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Rural</td>
</tr>
<tr>
<td>Target Tract</td>
<td>1045</td>
<td>1724</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>395</td>
<td>639</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>278</td>
<td>431</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>188</td>
<td>250</td>
</tr>
<tr>
<td>1 PUMA</td>
<td>164</td>
<td>207</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>143</td>
<td>182</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>107</td>
<td>116</td>
</tr>
<tr>
<td>1 State Away</td>
<td>89</td>
<td>101</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>85</td>
<td>95</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the distance bins captured by the row labels. The first five columns provide the estimated average job-related welfare gain (scaled to be equivalent to $ of 2010 annual earnings) from a 500 job stimulus package for workers whose distance between their origin jobs and the census tract receiving the stimulus package place them in the distance bin indicated in the row label. The next five columns provide the share of total stimulus-driven welfare gains that accrue to workers whose distance between their origin jobs and the census tract receiving the stimulus package place them in the distance bin indicated in the row label. Each column displays the average welfare outcome by distance bin among a subset of simulations featuring focal census tracts whose characteristics align with the column label. “All”: An average of all 500 target census tracts chosen as sites of simulated stimulus packages. “Rural”/“Urban”: An average over the 100 census tracts featuring the lowest/highest residential density (residents per square mile) among the full 500 target tracts simulated. “Small”/“Large”: An average over the 100 census tracts featuring the smallest/largest initial employment levels (based on total employment at establishments located in the tract) among the full 500 target tracts simulated.
Table A16: Heterogeneity in Change in P(Employed) and Share of Total Employment Gains by Distance from Focal Tract Across Focal Tracts of Varying Population and Employment Size

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>Change in P(Employed)</th>
<th>Share of Employment Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Rural</td>
</tr>
<tr>
<td>Target Tract</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>9.4E-04</td>
<td></td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>6.0E-04</td>
<td></td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>3.5E-04</td>
<td></td>
</tr>
<tr>
<td>1 PUMA</td>
<td>2.6E-04</td>
<td></td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>1.8E-04</td>
<td></td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>5.7E-05</td>
<td></td>
</tr>
<tr>
<td>1 State Away</td>
<td>2.0E-05</td>
<td></td>
</tr>
<tr>
<td>2+ States Away</td>
<td>6.7E-06</td>
<td></td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the distance bins captured by the row labels. The first five columns provide the estimated change in the probability of employment in the destination year caused by a 500 job stimulus package for workers whose distance between their origin jobs and the census tract receiving the stimulus package place them in the distance bin indicated in the row label. The next five columns provide the share of total stimulus-driven employment gains that accrue to workers whose distance between their origin jobs and the census tract receiving the stimulus package place them in the distance bin indicated in the row label. Each column displays the average welfare outcome by distance bin among a subset of simulations featuring focal census tracts whose characteristics align with the column label. “All”: An average of all 500 target census tracts chosen as sites of simulated stimulus packages. “Rural”/“Urban”: An average over the 100 census tracts featuring the lowest/highest residential density (residents per square mile) among the full 500 target tracts simulated. “Small”/“Large”: An average over the 100 census tracts featuring the smallest/largest initial employment levels (based on total employment at establishments located in the tract) among the full 500 target tracts simulated.
Table A17: Assessing the Value of Restricting Stimulus Jobs to Fill Positions Within the Target PUMA: Spatial Employment and Welfare Incidence for Restricted and Unrestricted Stimulus Packages (Each Featuring 500 Positions at a Large Low-Paying Manufacturing Firm)

<table>
<thead>
<tr>
<th>Distance from Target Tract</th>
<th>Change in P(Employed)</th>
<th>Share of Emp. Gains</th>
<th>Avg. Welfare Change ($)</th>
<th>Share of Wel. Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Tract</td>
<td>3.0E-03</td>
<td>1.4E-02</td>
<td>0.006</td>
<td>0.029</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>1.2E-03</td>
<td>4.0E-03</td>
<td>0.015</td>
<td>0.049</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>7.5E-04</td>
<td>2.1E-03</td>
<td>0.021</td>
<td>0.059</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>4.0E-04</td>
<td>7.3E-04</td>
<td>0.053</td>
<td>0.098</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>2.8E-04</td>
<td>2.0E-04</td>
<td>0.050</td>
<td>0.035</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>1.9E-04</td>
<td>1.4E-04</td>
<td>0.102</td>
<td>0.074</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>5.8E-05</td>
<td>4.5E-05</td>
<td>0.443</td>
<td>0.346</td>
</tr>
<tr>
<td>1 State Away</td>
<td>1.8E-05</td>
<td>1.6E-05</td>
<td>0.117</td>
<td>0.103</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>6.0E-06</td>
<td>6.5E-06</td>
<td>0.194</td>
<td>0.210</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the row labels and the outcomes in the column labels. Table entries consist of various measures of incidence by worker initial distance from the target census tract from a stimulus package consisting of 500 new jobs at large (top quartile of employment), low-paying (2nd quartile of avg. worker pay) manufacturing establishments. Columns labeled “Res.” report results from specifications in which the new positions are constrained to be filled by workers initially working (or most recently working) in the same PUMA as the targeted tract, while columns labeled “Unres.” report results from specifications in which the new positions may be filled by any worker in the nation.
Table A18: Changes in Employment Probabilities and Welfare from Alternative Relocation Packages for a Randomly Chosen Individual at Different Distances from the Receiving Tract (500 Positions Either Added Via a Stimulus Package or Relocated among Large Low-Paying Manufacturing Firms or Small High-Paying Information Firms)

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>Change in P(Employed)</th>
<th>Avg. Welfare Change (in $)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Tract</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>7.8E-04</td>
<td>7.9E-04</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>4.8E-04</td>
<td>4.8E-04</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>2.9E-04</td>
<td>2.9E-04</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>2.2E-04</td>
<td>2.3E-04</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>1.5E-04</td>
<td>1.6E-04</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>5.1E-05</td>
<td>5.5E-05</td>
</tr>
<tr>
<td>1 State Away</td>
<td>1.7E-05</td>
<td>2.2E-05</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>-2.1E-05</td>
<td>7.4E-06</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the distance bins captured by the row labels. Table entries in the first four columns consist of changes in the probability of employment in the destination year by worker initial distance from the target census tract from either a stimulus or a relocation package consisting of 500 new jobs either at small (lowest quartile of employment), high-paying (top quartile of avg. worker pay) establishments in the information supersector (in columns 1-2) or large (top quartile of employment), low-paying (2nd quartile of avg. worker pay) manufacturing establishments (in columns 3-4. The final four columns report results for the same specifications, except the outcome is the average job-related welfare gain (scaled to be equivalent to $ of 2010 annual earnings). The odd columns (labeled “reloc.”) report results from specifications in which the new jobs are relocated from locations at least 3 states away, while the even columns (labeled “stim.”) report results from stimulus specifications in which the new positions are added to the focal tract without being removed elsewhere (in the model, they are removed from the “nonemployment” destination type).
Table A19: Assessing the Impact on Employment and Welfare Outcomes of a Natural Disaster Removing 25, 50 or 100% of Positions in the Focal Tract for a Randomly Chosen Individual at Different Distances from Focal Tract Across (Averaging Across the Initial Earnings Distribution)

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>Change in P(Unemployed)</th>
<th>Share of Emp. Loss</th>
<th>Change in Welfare ($)</th>
<th>Share of Wel. Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25%</td>
<td>50%</td>
<td>100%</td>
<td>25%</td>
</tr>
<tr>
<td>Target Tract</td>
<td>0.029</td>
<td>0.073</td>
<td>0.192</td>
<td>0.126</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>4.4E-04</td>
<td>8.1E-04</td>
<td>1.4E-03</td>
<td>0.011</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>2.5E-04</td>
<td>4.7E-04</td>
<td>8.0E-04</td>
<td>0.015</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>1.5E-04</td>
<td>2.8E-04</td>
<td>5.0E-04</td>
<td>0.042</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>1.0E-04</td>
<td>2.0E-04</td>
<td>3.8E-04</td>
<td>0.037</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>6.6E-05</td>
<td>1.3E-04</td>
<td>2.6E-04</td>
<td>0.073</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>2.4E-05</td>
<td>4.8E-05</td>
<td>9.3E-05</td>
<td>0.386</td>
</tr>
<tr>
<td>1 State Away</td>
<td>8.4E-06</td>
<td>1.7E-05</td>
<td>3.2E-05</td>
<td>0.111</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>3.6E-06</td>
<td>5.6E-06</td>
<td>1.1E-05</td>
<td>0.199</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the row labels. The column labeled “Change in P(Unemployed)” indicates the change in the probability that a randomly chosen worker in the row subgroup will be unemployed in the destination year as a consequence of the simulated natural disaster. The column labeled “Change in Welfare” indicates the change in job-related welfare (scaled to be equivalent to $ of 2010 annual earnings) that a randomly chosen worker in the subgroup indicated by the row label will experience as a consequence of the simulated natural disaster. The columns labeled “Share of Emp. Loss” and “Share of Wel. Loss” indicate the share of all employment and welfare losses, respectively, generated by the simulated natural disaster that accrue to workers in the distance bin indicated by the row label. The column subheadings “25%”, “50%”, “100%” indicate the share of jobs in the focal tract that were removed in the simulations whose incidence is summarized in the chosen column.
Table A20: Share of Additional Unemployment and Welfare Losses Produced by a Natural Disaster Removing 25, 50 or 100% of Positions in the Focal Tract Among Workers at Different Initial Earnings Quartiles (or Unemployed)

<table>
<thead>
<tr>
<th>Earnings Quintile</th>
<th>Share of Emp. Loss</th>
<th>Share of Wel. Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25%</td>
<td>50%</td>
</tr>
<tr>
<td>UE ≤ Age 25</td>
<td>0.081</td>
<td>0.077</td>
</tr>
<tr>
<td>UE &gt; Age 25</td>
<td>0.267</td>
<td>0.254</td>
</tr>
<tr>
<td>1st Quartile</td>
<td>0.278</td>
<td>0.274</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>0.172</td>
<td>0.176</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>0.112</td>
<td>0.120</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>0.089</td>
<td>0.098</td>
</tr>
</tbody>
</table>

Notes: See Table A10 for expanded definitions of the origin employment status/earnings quartiles indicated by the row labels. The entries in the columns labeled “Share of Emp. Loss” and “Share of Wel. Loss” indicate the share of all employment and welfare losses, respectively, generated by the simulated natural disaster that accrue to workers in the initial employment status bin indicated by the row label. The column subheadings “25%”, “50%”, “100%” indicate the share of jobs in the focal tract that were removed in the simulations whose incidence is summarized in the chosen column.
Table A21: Change in Probability of Unemployment From a Natural Disaster Destroying either 25% or 100% of Positions in the Focal Tract Among Workers Initially Employed at Different Combinations of Initial Earnings Quartile (or Unemployed) and Distance from Focal Tract

| Distance from Focal Tract | 25% of Jobs Destroyed | | | | | | 100% of Jobs Destroyed | | | | |
|--------------------------|-----------------------|----------------|----------------|----------------|-----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Target Tract             | 0.002     | 0.006  | 0.044  | 0.036  | 0.028  | 0.022 | 0.006   | 0.016  | 0.218  | 0.233 | 0.225 | 0.194    |
| 1 Tct Away               | 7.4E-04   | 1.5E-03| 5.8E-04| 3.7E-04| 2.3E-04| 1.4E-04| 2.2E-03 | 4.3E-03| 2.0E-03| 1.3E-03| 8.9E-04| 4.8E-04 |
| 2 Tcts Away             | 4.9E-04   | 8.5E-04| 3.6E-04| 2.1E-04| 1.3E-04| 7.9E-05| 1.3E-03 | 2.5E-03| 1.2E-03| 6.9E-04| 4.2E-04| 2.5E-04 |
| 3+ Tcts w/in PUMA       | 2.5E-04   | 5.0E-04| 2.1E-04| 1.3E-04| 7.6E-05| 5.3E-05| 8.1E-04 | 1.6E-03| 7.3E-04| 4.4E-04| 2.7E-04| 1.6E-04 |
| 1 PUMA Away             | 1.8E-04   | 3.4E-04| 1.4E-04| 8.4E-05| 5.1E-05| 3.0E-05| 6.1E-04 | 1.2E-03| 5.4E-04| 3.2E-04| 2.0E-04| 1.2E-04 |
| 2 PUMAs Away            | 1.2E-04   | 2.3E-04| 9.1E-05| 5.4E-05| 3.2E-05| 2.0E-05| 4.4E-04 | 8.2E-04| 3.6E-04| 2.2E-04| 1.3E-04| 8.3E-05 |
| 3+ PUMAs w/in State     | 5.0E-05   | 8.7E-05| 3.2E-05| 2.0E-05| 1.2E-05| 6.8E-06| 1.8E-04 | 3.2E-04| 1.2E-04| 7.7E-05| 4.7E-05| 2.7E-05 |
| 1 State Away            | 1.7E-05   | 3.4E-05| 1.1E-05| 6.3E-06| 3.8E-06| 2.6E-06| 6.3E-05 | 1.3E-04| 4.3E-05| 2.4E-05| 1.5E-05| 1.0E-05 |
| 2+ States Away          | 6.4E-06   | 1.3E-05| 3.9E-06| 2.1E-06| 1.2E-06| 8.9E-07| 2.5E-05 | 4.9E-05| 1.5E-05| 8.2E-06| 4.8E-06| 3.4E-06 |

Notes: See Table A4 for expanded definitions of the distance bins represented by the row labels. See Table A10 for expanded definitions of the origin employment status/earnings quartiles indicated by the column labels. Each entry provides the average increase in the probability of unemployment from simulations in which either 25% or 100% of the initial jobs in the chosen census tract are removed and replaced with “unemployment” positions for workers whose initial job (or most recent job if initially unemployed) is located in the distance bin associated with the row label, and whose initial employment status or earnings quartile (if initially employed) falls into the employment status bin associated with the column label. The average is taken across 500 simulations featuring different target census tracts.
Table A22: Expected Change in Utility From a Natural Disaster Removing either 25% or 100% of Positions in the Focal Tract Among Workers Initially Employed at Different Combinations of Initial Earnings Quartile (or Unemployed) and Distance from Focal Tract

<table>
<thead>
<tr>
<th>Focal Tract</th>
<th>25% of Jobs Destroyed</th>
<th>100% of Jobs Destroyed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Tract</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>-142</td>
<td>-134</td>
</tr>
<tr>
<td>1 State Away</td>
<td>-118</td>
<td>-118</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the distance bins represented by the row labels. See Table A10 for expanded definitions of the origin employment status/earnings quartiles indicated by the column labels. Each entry provides the average increase in job-related welfare (scaled to be equivalent to 8 of 2010 annual earnings) from simulations in which either 25% or 100% of the initial jobs in the chosen census tract are removed and replaced with “unemployment” positions for workers whose initial job (or most recent job if initially unemployed) is located in the distance bin associated with the row label, and whose initial employment status or earnings quartile (if initially employed) falls into the employment status bin associated with the column label. The average is taken across 500 simulations featuring different target census tracts.
Table A23: Change in Probability of Destination Employment (or Unemployment) at Different Distances from Focal Tract after a Natural Disaster Removing 25%, 50% or 100% of Positions for Workers Initially Employed (or Most Recently Employed) in the Focal Tract (Averaging Across the Initial Earnings Distribution)

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>% of Jobs Removed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25%</td>
</tr>
<tr>
<td>Unemployment</td>
<td>0.029</td>
</tr>
<tr>
<td>Target Tract</td>
<td>-0.106</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>0.005</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>0.005</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>0.009</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>0.008</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>0.012</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>0.022</td>
</tr>
<tr>
<td>1 State Away</td>
<td>0.003</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the distance bins represented by the row labels. Each entry gives the change in the probability of employment at a location whose distance falls into the distance bin associated with the row label for workers initially working (or most recently working) in the focal census tract for a simulated natural disaster in which 25%, 50% or 100% of jobs are removed (indicated by the column label). Each entry represents an average over 500 simulations featuring different target census tracts. The entries in the row labeled “Unemployment” provides the change in the share of workers who stay or become unemployed caused by the natural disaster.
Table A24: Change in Probability of Destination Employment (or Nonemployment) at Different Distances from Focal Tract after a Natural Disaster Removing either 25% or 100% of Positions for Workers Initially Employed in the Focal Tract by Initial Earnings Quartile (or Nonemployment)

<table>
<thead>
<tr>
<th>Distance from Focal Tract</th>
<th>25% of Jobs Destroyed</th>
<th>100% of Jobs Destroyed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonemployment</td>
<td>0.002</td>
<td>0.006</td>
</tr>
<tr>
<td>Target Tract</td>
<td>-0.010</td>
<td>-0.010</td>
</tr>
<tr>
<td>1 Tct Away</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>2 Tcts Away</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>3+ Tcts w/in PUMA</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>1 PUMA Away</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>2 PUMAs Away</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>3+ PUMAs w/in State</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>1 State Away</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>2+ States Away</td>
<td>0.001</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Notes: See Table A4 for expanded definitions of the distance bins represented by the row labels. See Table A10 for expanded definitions of the origin employment status/earnings quartiles indicated by the column sublabels. Each entry gives the change in the probability of employment at a location whose distance falls into the distance bin associated with the row label and whose initial employment status/earnings quartile falls into the bin associated with the column sublabel for workers initially working (or most recently working) in the focal census tract due to a simulated natural disaster in which either 25% or 100% of jobs are removed. Each entry represents an average over 500 simulations featuring different target census tracts. The entries in the row labeled “Unemployment” provides the change in the share of workers who stay or become unemployed caused by the natural disaster.
Table A25: Model Validation Results: Dissimilarity Index Values Comparing Forecasted and Actual Worker Reallocations Following Large Local Shocks Using Alternative Transition Group Definitions and Methods for Generating Forecasts

<table>
<thead>
<tr>
<th>Level of Group Aggregation</th>
<th>All U.S.</th>
<th>Target PUMA Only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Two-Sided Matching</td>
<td>Param. Logit</td>
</tr>
<tr>
<td>Full Group Space</td>
<td>0.071 (0.001)</td>
<td>0.427 (0.002)</td>
</tr>
<tr>
<td>Sm. Dist. Bins</td>
<td>0.053 (0.001)</td>
<td>0.419 (0.002)</td>
</tr>
<tr>
<td>Sm. Dist. Bins &amp; No Firm Char.</td>
<td>0.009 (0.000)</td>
<td>0.195 (0.003)</td>
</tr>
<tr>
<td>Lg. Dist. Bins &amp; No Firm Char.</td>
<td>0.007 (0.000)</td>
<td>0.184 (0.002)</td>
</tr>
<tr>
<td>Unemp. Only (All Loc.)</td>
<td>0.059 (0.001)</td>
<td>0.148 (0.003)</td>
</tr>
<tr>
<td>Unemp. Only (Lg. Dist. Bins)</td>
<td>0.043 (0.001)</td>
<td>0.125 (0.003)</td>
</tr>
<tr>
<td>E-to-UE and UE-to-E Only (All Loc.)</td>
<td>0.008 (0.000)</td>
<td>0.333 (0.006)</td>
</tr>
<tr>
<td>E-to-UE and UE-to-E Only (Lg. Dist. Bins)</td>
<td>0.044 (0.002)</td>
<td>0.715 (0.007)</td>
</tr>
</tbody>
</table>

Notes: This table examines the fit of model-based predicted worker reallocations to the actual reallocations that occurred following a set of local employment shocks to particular census tracts in particular years between 1996-2010. See Section A7 for a detailed description of the model validation exercise. Each row of the table considers a different metric for measuring model fit, while each column considers a different combination of model and target population. Columns 1-4 examine the job reallocation fit among all U.S. citizens in my 19 state LEHD sample, while columns 5-8 consider the fit only among workers initially working in the same PUMA as the tract receiving the shock. Each entry averages the fit metric across all 514 local shocks identified. For each shock, predictions are based on parameters estimated using local data from the year before the shock occurred. “Two-sided Matching” refers to the preferred two-sided matching model presented in this paper. “Param. Logit” refers to a one-sided parametric conditional logit model (See A7 for a list of the predictor variables). “Raw CCP” refers to a prediction that holds the previous year’s conditional choice probability (CCP) distribution constant for each destination type, but updates the destination type marginal distribution to reflect the shock, while “Smoothed CCP” does the same but smooths the CCPs across similar destination types before constructing the predicted reallocation. None of the three alternative models impose market clearing. “Full Group Space” evaluates model fit using the index of dissimilarity between the actual and predicted distribution across groups in the transition group space. “Sm. Dist. Bins”, “Sm. Dist. Bins & No Firm Char” and “Lg. Dist. Bins & No Firm Char” evaluate the index of dissimilarity on aggregated group spaces in which origin and destination locations are each aggregated to small or large distance bins relative to the focal tract, and, in the latter two cases, destination types featuring the same distance bin but different non-location characteristics are combined. “Unemp. Only (All Loc.)” evaluates the index of dissimilarity between predicted and actual shares of unemployed workers after the shock originally located in each origin location. “Unemp. Only (Lg. Dist. Bins)” does the same but aggregates origin locations to coarse distance bins relative to the focal census tract. “E-to-UE and UE-to-E Only (All Loc.)” calculates the index of dissimilarity only among transition groups featuring employment-to-unemployment and unemployment-to-employment transitions, while “E-to-UE and UE-to-E Only (Lg. Dist. Bins)” does the same but aggregates locations during employment to large distance bins relative to the focal census tract.