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

June 14, 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 demand 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 those 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 early assistance from Evan Buntrock, John Abowd, and Erika McEntarfer and valuable comments from Jeronimo Carballo, Terra McKinnish, Brian Cadena, Francisca Antman, Miles Kimball, Martin Boileau, Alessandro Peri, Murat Iyigun, Stephan Heblich, Yanos Zylberberg, and participants at the NBER Trade and Geography Conference, the IZA World of Labor Conference, the American and European Urban Economics Association Conferences, and seminars at the University of Illinois, Princeton University, Colorado State University, the University of Bristol, and 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 of various skill classes 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 employment and welfare incidence expected to redound to workers within the city, county, and state borders, respectively. Officials need to be able to predict which types of targeted firms will yield a geographically concentrated impact in which the labor demand shock primarily trickles down to lower skill levels rather than out toward 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. More distant towns, counties or states are 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 single 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 U.S. local labor markets. These simulations 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 (hereafter CS) 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. 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 groups defined by races, age groups, or industries (or combinations thereof).

Second, the assignment game requires market clearing, optimizing behavior by all 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 model parameters (mean relative joint surpluses among matched pairs of workers and firms belonging to observable types) can be identified from a single labor market transition from origin to destination job assignments, and are sufficient to perform counterfactuals that yield the resulting allocation and payoff changes for all players (workers and firms) from any arbitrary change in the composition of labor supply and/or demand.

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 to either the type composition of labor supply and/or demand or to composite joint surplus parameters. Furthermore, while 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 model predictions 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, policymakers concerned that local initiatives creating new high-skilled positions could increase rent for low-skilled renters might wish to know whether their downstream earnings and employment opportunities 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 jurisdictions both near and far from the shock.

The counterfactual simulations involve stimulus projects that create 500 new job positions in particular U.S. locations (census tracts) featuring alternative combinations of establishment size, average pay, and industry supersector as well as "natural disaster" simulations akin to a tornado or flood that eliminate a share of all jobs in a location.

I find that job-related welfare gains or losses from very local shocks are widely distributed. For example, the simulations suggest that as little as 7.0% of the utility gains and 8.1% of the net employment gains from such local stimuli 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 even though the same simulations suggest that a randomly chosen worker in the targeted tract is about 3, 27 and over 9,000 times more likely to fill one of the new vacancies than counterparts in an adjacent tract, an adjacent PUMA, and a non-adjacent state, respectively. These seemingly inconsistent findings are the result of two mechanisms.

First, workers who join the incoming firms are likely to already be employed, so their transitions generate further openings for others. Thus, 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 importantly), because a single census tract generally only contains a few thousand workers, its workforce makes up a very small share of the national labor market: 73,057 tracts and 2,378 PUMAs were defined in the 2010 U.S. census. Consequently, even quite disproportionate welfare gains for the most local workers cannot account for more than a tiny share of the aggregate welfare gains. Indeed, the mean predicted utility and employment gains for initially local workers are 2.6 (3.2), 3.8 (8), and 12.3 (448) times as large as for workers in an adjacent tract, an adjacent PUMA, and 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 nationally), 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 $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 (equivalent to $1242) and prime-age non-employed ($1165), and smallest among initially low-paid workers ($999) and the young non-

[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.

employed ($620), where I use prior earnings as a proxy for worker skill. However, these averages mask substantial heterogeneity in projected impacts across stimuli featuring different establishment compositions or locations. The average utility gains for the same four skill/age groups are $1728, $1015, $825 and $415 for those 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 establishment composition, as the simulated shocks ripple outward, they become less skill-biased: predicted differences in welfare (or employment) gains among skill groups converge as one considers 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 ($\approx$15% vs. 7.5%), and that requiring the newly created jobs to be filled exclusively from existing PUMA workers (or jobseekers) increases their share of employment gains to nearly 25% from under 10%.

To capture such multidimensional heterogeneity, the surplus parameters that determine equilibrium elasticities of substitution (and thus govern the impact of simulated shocks) exploit for identification the full sample of worker flows generated by an implicit mix of local and national shocks to both supply and demand composition, rather than exclusively using flows involving locations experiencing local labor demand 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 flows from prior years are used to predict the realized reallocation around 514 census tracts that experienced gains or losses of between 100 and 3000 jobs within one year between 1996 and 2010. The model predicts these out-of-sample reallocations quite well and considerably better than even relatively rich one-sided parametric models that fit firm or worker conditional choice probabilities with over 100 parameters. This validation exercise illustrates that the very large set of estimated parameters is not causing overfitting, but is instead necessary to capture the highly nonlinear and multidimensional nature of the U.S. job matching technology.

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).[2]

Autor et al. (2014)'s evaluation of the worker-level impact of China's accession to the WTO is

[2] A 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).

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 commuting adjustments may largely facilitate the shock response, limiting the impact on rent.[3]

A related sub-literature seeks to estimate local job multipliers due to the increased product demand and agglomeration and congestion externalities created by an initial infusion of new positions (e.g. Moretti (2010) or Bartik and Sotherland (2019)). My approach complements this sub-literature: those papers are generally silent about which types of workers from which initial locations benefit most from the estimated net change in local job opportunities, while the assignment model here takes the new spatial distribution of positions (possibly including jobs generated by multipliers) as an exogenous input and evaluates the resulting skill and spatial incidence.

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. Schmutz and Sidibe (2016) estimate a search-and-matching model with data on worker flows among French metropolitan areas. They show that search frictions play a greater role than moving costs in limiting worker mobility and determining the incidence of local shocks, suggesting the potential promise of efforts to disseminate information about distant jobs. Monte et al. (2015) and Caliendo et al. (2015) (hereafter CDP) each estimate trade-theoretic models with labor, housing, and product market clearing and arbitrary spatial frictions in both labor and product markets. The former features joint choices of residential and work locations, and highlights the role of commuting in determining local shock incidence.[4] The latter shows that counterfactual dynamic equilibrium paths can be evaluated for alternative structural shocks (changes in trade costs, mobility costs, productivities, etc.) without estimating all the model primitives. 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 any of the fundamental utility, production function, and moving cost parameters. 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.

Each of these papers aggregates locations to at least the county level. Manning and Petrongolo (2017), by contrast, represents the most notable attempt to determine the equilibrium incidence

[3]The 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).

[4]Due to a lack of residential microdata, I do not consider whether new job matches involve residential mobility.

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 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 dynamics and 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 worker heterogeneity beyond initial location, and only CDP (industry differences) features any observable firm heterogeneity besides firm location. Similarly, several spatial equilibrium models in the labor literature (e.g. Piyapromdee (2017) or Diamond (2016)) feature imperfect substitution among observable worker types, but only differentiate firms by 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. While Lindenlaub (2017) and Bonhomme et al. (2019) each estimate multidimensional labor market sorting models, they consider assignment of workers to occupations and jobs without incorporating geography or spatial frictions.

Indeed, Nimczik (2017), who characterizes labor markets as networks of firms disproportionately sharing worker flows, shows that the geographic and industrial scope of labor markets varies substantially across occupation and education categories. However, the stochastic block model he employs separately defines labor markets for each skill category, so that the various skill categories cannot be incorporated into a single integrated labor market. Thus, unlike an equilibrium model with two-sided optimal choice, it is not designed to analyze the tradeoffs firms and workers make following local demand shocks between settling for skill mismatch and paying the moving and search costs needed to overcome spatial mismatch.

Finally, this paper also 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 empirical application of a two-sided assignment model to the national labor market.[5] The theoretical properties of assignment games have been well-established for at least a generation.[6] However, the empirical content of assignment games and hedonic models 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 pioneering work by Ekeland et al. (2004), Heckman et

[5] See Tervio (2008) and Chen (2017) for applications of the assignment game to the narrower market for CEOs.

[6] See Koopmans and Beckmann (1957), Shapley and Shubik (1972), Roth and Sotomayor (1992), and Sattinger (1993).

al. (2010) and CS leading to contributions by Chiappori and Salanié (2016), Menzel (2015), and Galichon and Salanié (2015), among others. I make two contributions to this literature.

First, I consider implementation in a context featuring millions of match observations and thousands of types on both the supply and demand side. I address the challenge of a somewhat sparse matching matrix by developing a smoothing procedure 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 when the number of unmatched partners of each type is either unobserved or only observed on one side of the market: while nonemployment by type may be inferred with reasonable accuracy by combining LEHD and public-use ACS data, unfilled vacancy counts by detailed type are not available. The existing identification results of CS and Menzel (2015), among others, rely on observing the number of singles on both sides. I discuss conditions under which ignoring unmatched partners would not affect the incidence of local shocks.

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 applies the insights of CS 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. In addition, Section 5 also introduces the various labor demand shocks 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.

2 The Two-Sided Assignment Model

In this section I model the evolution of the labor market across adjacent time periods as a static assignment game played by workers and establishments. The model introduces several features and extensions necessary to adapt CS's marriage market model to a labor market setting. The exposition closely mirrors Galichon and Salanié (2015) (hereafter GS), which generalizes CS. Section 2.1 defines the matching game. Section 2.2 describes how the workers and positions and the job matches that determine the game's payoffs are aggregated to types and groups, respectively. Section 3.1 imposes additional structure 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 to use the estimated surpluses to construct counterfactual simulations capturing the incidence of local labor supply and demand shocks of varying worker and position compositions.

2.1 Defining the Assignment Game

Suppose that in year y there are I potential workers comprising the set $\mathcal{I}$ who participate in the labor market. Each worker i begins the year in a job match, determined in year $y-1$, with a position

$j(i)$ at establishment $m(j(i))$ taken from the set of possible positions $\mathcal{J}$. Let $m(j) = 0$ 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 $U(i, j(i), k)$, or more parsimoniously as $U(i, k)$. The worker's potential annual earnings in year y from accepting position k, denoted w_{ik}, is assumed to be additively separable from all other determinants of the worker's payoff, so that $U(i, k)$ takes on a quasi-linear money-metric form:[7]

$$U(i, k) = \pi^i_{ik} + w_{ik} \tag{1}$$

π^i_{ik} 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 components or the functions governing their links to payoffs to construct the counterfactual simulations that form the primary contribution of the paper. Any value function specification in which current worker earnings are additively separable will suffice. That said, careful thought about which determinants of the payoff are likely to be large and differential across alternative workers, positions, and job matches is necessary to guide the choice of characteristics used to assign workers and positions to types in section 2.2 below, as well as to evaluate the plausibility of assumptions laid out in section 3.2 that underlie the simulations.

Such components might include worker i's valuation of various non-pecuniary amenities offered by position k (including the appeal of its geographic location), as well as any search, moving, or training costs paid by worker i to find, move to, or settle into position k from initial position j.[8] They might also include the continuation value associated with beginning year $y + 1$ 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.

On the other side of the market there are K potential positions comprising set $\mathcal{K}$ at establishments that seek workers in year y. The intersection of $\mathcal{K}$ and $\mathcal{J}$ may be quite large, so that many end-of-period positions in $\mathcal{K}$ can potentially be "filled" by continuing a job match that already exists. I assume that each establishment makes independent hiring decisions for each position, so as to model positions' over individual workers rather establishments' preferences over collections of workers.[9] Let $V(i, k)$ denote the value to position k in establishment $m(k)$ of hiring (or retaining) worker i. The potential annual earnings paid by k to worker i in year y is assumed to be additively

[7]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.

[8]The traditional assignment game precludes stochastic search frictions, so that each agent may potentially match with any agent on the opposite side. However, Menzel (2015) shows that one can introduce a probability that i and k meet that is independent of other payoff determinants, assign their joint surplus 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 to an intermediary (e.g. a headhunter or a matching website) to reveal certain opposite side agents.

[9]One justification for treating positions as independent is that there are nontrivial costs of coordinating multiple independent hires/retentions that outweigh the gains from better exploiting production complementarities. Roth and Sotomayor (1992) highlight the complications that arise when establishments have preferences over collections of workers.

separable from all other determinants of the position's payoff, so that $V(i,k)$ can be written as:

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

Akin to π_{ik}^{i}, π_{ik}^{k} combines the contributions of several payoff components that need not be fully specified by a particular functional form. These components might include worker i's contribution to establishment $m(k)$'s revenue in year y, any recruiting, moving, and training costs borne by $m(k)$ in hiring worker i, as well as any continuation value from starting year $y+1$ with i already installed in position k, including the fact that retaining i next year avoids 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:

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

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. Written in this form, the game's structure mimics the classic assignment game analyzed by Shapley and Shubik (1972).

A matching or market-wide transition in this labor market is an $I \times K$ transition matrix μ such that $\mu_{i,k}$ = 1 if worker i matches with position k, and 0 otherwise. As in Galichon and Salanié (2015), I focus on stable matchings, which require a division of joint surplus 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 under the proposed matching. Shapley and Shubik (1972) show that the set of stable matchings coincides with both the core of the game and with the set of competitive equilibria from a decentralized market. Furthermore, they show that with 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 maximizes the aggregate surplus and is the solution to a linear programming problem.[10]

Equivalently, the unique stable assignment can also be found by solving the dual problem: identifying a set of worker utility values $\{r_i\}$ and position profit values $\{q_k\}$ that minimize the total "cost" of all workers and positions, $\sum_{i\in\mathcal{I}} r_i + \sum_{k\in\mathcal{K}} q_k$, subject to the constraint that these values cannot violate the underlying joint surplus values: $r_i + q_k \geq \pi_{ik} \; \forall \, (i,k)$. Crucially, inspection of the problem reveals that the stable assignment is fully determined by the joint surplus values $\{\pi_{ik}\}$; no separate information on the worker and firm components π_{ik}^{i} and π_{ik}^{k} is needed. This dual problem yields the following conditions that define the optimal assignment (GS):

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

[10] The joint surplus is given by $\sum_{(i,k)\in\mathcal{I}\times\mathcal{K}} \mu_{i,k}\pi_{ik} + \sum_{i\in\mathcal{I}:\mu_{i,k}=0 \; \forall \; k} \mu_{i,0}\pi_{i0} + \sum_{k\in\mathcal{K}:\mu_{i,k}=0 \; \forall \; i} \mu_{0,k}\pi_{0k}$, where π_{i0} and π_{0k} denote i's payoff from unemployment and k's payoff from remaining vacant. Constraints must also hold that each position and worker can match with at most one counterpart: $\sum_i \mu_{i,k} \leq 1 \; \forall \; k \in \mathcal{K}$ and $\sum_k \mu_{i,k} \leq 1 \; \forall \; i \in \mathcal{I}$.

I aggregate these conditions in the next section to deliver identification of "group"-level surpluses. Given optimal 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}:

$$w_{ik} = \pi_{ik}^k - q_k \tag{5}$$

Because $r_i + q_k = \pi_{ik} \equiv \pi_{ik}^i + \pi_{ik}^k$ for any matched pair (i, k) in the stable match, this implies:

$$w_{ik} = r_i - \pi_{ik}^i \tag{6}$$

Using (5) and (6), the conditions (4) can be rewritten as the standard requirements that worker and establishment choices must be utility- and profit-maximizing, respectively:

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

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 will be exploited below. Importantly, while the stable assignment μ is generally unique, the equilibrium payoffs and transfers are not.[11] The exact equilibrium payoffs/wages that emerge depend on the particular market clearing mechanism.

While the model does not require a particular earnings determination process, one candidate is a simultaneous ascending auction in which all positions bid on all workers. Workers set reservation utilities based on their values of remaining unemployed for a year. Each position bids utility values of a one period commitment U_{ik} (which include the value of beginning the next period as an incumbent), and may only win the bidding for a single worker (or it may choose to remain vacant). The position k that bids the highest utility r_i retains worker i and pays an annual earnings amount w_{ik} that, combined with the non-pecuniary component π_{ik}^i, equals the worker's promised valuation $U_{ik} = r_i$. The auction ends when no position wishes to change its bid for any worker. Some workers may remain unemployed, and some positions may remain unfilled. Importantly, though positions start at different π_{ik}^i baselines, with transferable utility bid changes can always take the form of annual earnings increases. Thus, one may scale the changes in equilibrium utilities r_i following labor demand shocks in terms of annual earnings gains (though sometimes workers achieve utility gains by taking an earnings cut to join an establishment offering superior non-pecuniary values).

2.2 Modeling the Match Surpluses

Part of the joint transition surplus π_{ik} from transition (i, k) will be common to any transition 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 recruit-

[11] All $\{r_i\}$ values can generally be shifted slightly up or down (with offsetting $\{q_k\}$ shits) without violating stability.

ing 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 $g \in \mathcal{G}$ (with $G \equiv |\mathcal{G}|$), and use the notation $g(i, k)$ to denote the group assignment of transition (i, k). Importantly, these groups are always defined by a combination of *observable* characteristics of the worker and accompanying origin job match, the destination position, or the full transition (i, k). These characteristics should be chosen to capture as comprehensively but parsimoniously as possible the underlying (structural) preferences, productivities, moving costs, and search costs that determine the joint surpluses $\{\pi_{ik}\}$.

Some of the observed characteristics may only relate to the worker i, the origin position $j(i)$, or the origin job match $(i, j(i))$, and will be common to all destination positions. I use this subset of characteristics to assign each origin 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 this assignment. In the empirical work, the origin types (described in detail below) are defined by unique combinations of the origin establishment's geographic location and a prior earnings/unemployment category (used as a proxy for worker skill).[12]

Analogously, a second subset of the characteristics defining transition groups may only describe the destination position, and may be used to assign each destination position $k \in \mathcal{K}$ to a destination type $d \in \mathcal{D}$, denoted $d(k)$. In the empirical work, destination types consist of unique combinations of the following establishment characteristics: $m(k)$'s geographic location (detailed further below), its year $y - 1$ quartiles in the national establishment-level employment and average earnings distributions (the latter used to proxy for average skill requirements), and its industry supersector.[13]

Finally, let $z(i, j(i), k) \equiv z(i, k)$ denote any remaining characteristics defining the transition group that depend on both $(i, j(i))$ and k. In the empirical work below, the only z characteristic is an indicator for whether the "transition" represents continued employment at the same establishment, $1(m(k) = m(j(i)))$, so that job stayers and job movers are placed into different groups, reflecting the fact that search, recruiting, and training costs do not have to be repaid for existing workers. This 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).[14]

Thus, one can rewrite the mapping $g(i, k)$ as $g(o(i), d(k), z(i, k)) \equiv g(o, d, z)$. While knowledge of g is sufficient to recover o and d, knowledge of o and d need not uniquely identify 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 group g's origin type and $d(g) = d(g(i, k)) = d(k)$ to refer to its destination type.

Given these definitions, one can decompose the surplus π_{ik} into a part common to all transitions

[12] Ideally, residential location would define the origin job type and establishment location would define the destination job type. In the absence of data on workers' residential locations, origin establishment locations are used as proxies.

[13] Note that i's earnings in year $y-1$ is used to proxy for worker skill, but $m(k)$'s average pay in year y is used to proxy for k's skill requirements. This can be rationalized by assuming that a newly hired worker develops the required skills by the end of the year (perhaps incurring training costs affecting π_{ik} that are paid by either the worker or establishment).

[14] Mourifié and Siow (2017) use the same approach to distinguish marriage from cohabitation.

classified as group $g(i,k)$, denoted θ_g, and an idiosyncratic component ϵ_{ik} specific to (i,k):

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

ϵ_{ik} might reflect, for example, the low psychic costs of a particular worker who is moving back to a familiar location, or perhaps particular skill requirements of position k that worker i uniquely possesses. Following Decker et al. (2013), σ is a scaling parameter that captures the relative importance of idiosyncratic surplus components compared to components that are common to all transitions classified into group g in determining the variation in match surpluses across potential pairs $(i,k) \in \mathcal{I}\times\mathcal{K}$. I show below that counterfactual unique stable job assignments will not depend on σ, but σ plays a key role in determining the size of changes in offered utilities r_i for particular workers in particular locations necessary to facilitate the reallocation that yields the stable assignment.

The goal is to use the observed matching μ to recover the set of group mean surplus values $\{\theta_g\}$. As GS emphasize, one could impose further structure on the production, utility, and search cost, and recruiting cost functions that comprise the joint surpluses and estimate the model via maximum likelihood. Driven by computational considerations and an interest in being agnostic about the various structural functions that underlie $\{\theta_g\}$, I follow CS and leave the set $\{\theta_g\}$ unrestricted, achieving identification instead by assuming that the ϵ_{ik} draws are i.i.d across all alternative matches (i,k') and $(i',k) \in \mathcal{I}\times\mathcal{K}$ and follow a Type 1 extreme value distribution.[15] Unlike in CS and GS, equation (8) allows the idiosyncratic component to be truly pair-specific: the combined surplus from two matches changes if the two workers swap positions, even if the workers share an origin type and the positions share a destination type. Given coarsely-defined origin and destination types, pair-level heterogeneity in job match quality is likely to be substantial. But allowing such heterogeneity 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 θ_g into group-level worker and position subcomponents θ_g^l and θ_g^f analogous to the transition-level components π_{ik}^i and π_{ik}^k defined above. Fortunately, this decomposition is not necessary to generate key measures of worker-level incidence.

3 Identification

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

Recall from section 2.1 that a necessary condition for a matching μ to be stable (and thus sustainable as a competitive equilibrium) is that there exists a set of worker payoffs $\{r_i\}$ such that $\mu_{ik} = 1$ implies that $i \in \arg\max_{i\in\mathcal{I}} \pi_{ik} - r_i$ for any potential match $(i,k) \in \mathcal{I}\times\mathcal{K}$. Given candidate equilibrium payoffs $\{r_i\}$ combined with the i.i.d. Type 1 EV assumption for ϵ_{ik}, Decker et al.

[15]Menzel (2015) shows that imposing that ϵ_{ik} draws are i.i.d is the key assumption (the choice of the Type 1 extreme value distribution is not critical), since it causes each sides' conditional choice probabilities to satisfy the independence of irrelevant alternatives property. GS and Chiappori et al. (2009) discuss the possibility of allowing 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.

(2013) show that the probability that hiring (or retaining) i maximizes k's payoff is given by:

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

Next, define $f(o)$ as the share of workers assigned to type o, define C_o as the mean of $e^{-\frac{r_i}{\sigma}}$ among type o workers, and define $\overline{S}_{g|o,d}$ as the mean among type d positions of the share of type o workers whose hire/retention would be assigned to group g (i.e. the incumbent share if $z(g) = 1$, the mover share if $z(g) = 0$). With two additional assumptions, one can aggregate equation (9) to derive a tractable expression for the conditional probability $P(g|d)$ that a randomly chosen position of type d wishes to hire a worker whose transition would be assigned to group g:

$$P(g|d) = \frac{e^{\frac{\theta_g}{\sigma}} \overline{S}_{g|o,d} f(o) C_o}{\sum_{o' \in \mathcal{O}} \sum_{g' \in (o,d)} e^{\frac{\theta_{g'}}{\sigma}} \overline{S}_{g'|o',d} f(o') C_{o'}} \tag{10}$$

In particular, this expression depends only on the group g and the types o and d rather than the individual workers i and positions k.[16] Appendix A1 presents and proves this result formally as Proposition A1. Intuitively, the first assumption imposes that the utility payoffs required in equilibrium by workers from the same skill class and local area must not differ systematically across origin establishments. This becomes a better approximation as more characteristics (such as occupation or education) and finer geography are used to define origin types, so that workers of the same type become close substitutes for one another. The second assumption imposes that establishments of the same destination type feature roughly the same number and origin type distribution of incumbent workers. This approximation improves as narrower worker location and skill and establishment location, size and average pay categories are used to define origin and destination types.[17]

Next, let $\hat{\mu}$ denote an observed matching. Since each transition can be assigned to a unique group g, one can easily aggregate the matching into an empirical transition group distribution. Specifically, let $\hat{P}_g$ denote the fraction of observed transitions that are assigned to group g: $\hat{P}_g \equiv \frac{1}{|I|} \sum_{(i,k) \in \mathcal{I} \times \mathcal{K}} \hat{\mu}_{ik} 1(g(i,k) = g)$. Similarly, $\hat{f}(o)$ denotes the fraction of transitions featuring type o workers, $\hat{f}(o) = \frac{1}{|I|} \sum_{i \in \mathcal{I}} 1(o(i) = o)$, and $\hat{h}(d)$ denotes the fraction featuring type d destination positions, $\hat{h}(d) = \frac{1}{|K|} \sum_{k \in \mathcal{K}} 1(d(k) = d)$.[18] Given these definitions, one can estimate the conditional choice probability $P(g|d)$ by calculating the observed fraction of type d positions that were

[16]Note that in contrast to CS, the probability that a worker of origin type o is chosen depends on the share of workers of type o in the population, $f(o)$. This difference stems from allowing the idiosyncratic surplus component to be pair-specific. Menzel (2015) derives a similar formula in his nontransferable utility assignment model.

[17]These assumptions are necessary because the probability of filling a position with an existing employee depends on how many employees one already has, so that the group average of $e^{\frac{-r_i}{\sigma}}$, a non-linear function of the random variable r_i, depends on the entire establishment size distribution among firms who are at risk of creating a transition that could be classified into g. The assumptions essentially impose that Jensen's inequality is close to an equality for $e^{\frac{-r_i}{\sigma}}$.

[18]Note that in the empirical work each transition will have both an origin and a destination because I do not observe unfilled vacancies and I augment $\mathcal{K}$ to include a sufficient number of nonemployment "positions". As a result, the number of workers seeking jobs will equal the number of destination "positions".

filled via group g transitions: $\hat{P}(g|d) = \frac{\hat{P}_g}{\hat{h}(d)}$. As the number of observed transitions gets large, each member of the set of empirical CCPs $\{\hat{P}(g|d)\}$ should converge to the corresponding expression in (10). Note also that the average shares $\{\overline{S}_{g|o,d}\}$ can be estimated using the average across all possible transitions (i,k) assignable to group g of the incumbent indicator $1(m(j(i)) = m(k))$.

One may now assess the amount of information contained in the observed empirical choice probabilities $\{\hat{P}(g|d)\}$ about the mean match surplus values $\{\theta_g\}$. First, using (10), we can derive an expression for the log odds between two CCPs involving an (arbitrarily chosen) destination type d_1 and two (arbitrarily chosen) transition groups g_1 and g_2 for which $d(g_1) = d(g_2) = d_1$:

$$\ln(\frac{\hat{P}_{g_1|d_1}}{\hat{P}_{g_2|d_1}}) = (\frac{\theta_{g_1} - \theta_{g_2}}{\sigma}) + \ln(\frac{\overline{S}_{g_1|o(g_1),d_1}}{\overline{S}_{g_2|o(g_2),d_1}}) + \ln(\frac{f(o(g_1))}{f(o(g_2))}) + \ln(\frac{C_{o(g_1)}}{C_{o(g_2)}}) \tag{11}$$

Since the initial origin-type stocks $f(o(g_1))$ and $f(o(g_2))$ and shares of potential stayers $\overline{S}_{g_1|o(g_1),d_1}$ and $\overline{S}_{g_2|o(g_2),d_1}$ are either directly estimable or observed (depending on whether a sample or the full population is available), to establish identification one can treat their terms as known and bring them to the left hand side. Even these adjusted log odds still conflate the relative mean (re-scaled) surplus values from transition groups g_1 and g_2, $(\frac{\theta_{g_1} - \theta_{g_2}}{\sigma})$ with the log ratio of mean exponentiated worker re-scaled utilities between the two origin job types, $\ln(\frac{C_{o(g_1)}}{C_{o(g_2)}})$.

However, consider two additional groups g_3 and g_4 for which $d(g_3) = d(g_4) = d_2$ and $o(g_3) = o(g_1)$ and $o(g_4) = o(g_2)$.[19] The four groups g_1 to g_4 can be chosen to represent the two ways to match a given pair of destination positions to a given pair of workers. Taking the ratio of equation (11) to its analogue using g_3 and g_4 conditional on d_2 and rearranging, one obtains:

$$\ln(\frac{\hat{P}_{g_3|d_2}/(\overline{S}_{g_3|o(g_3),d_2} f(o(g_3)))}{\hat{P}_{g_4|d_2}/(\overline{S}_{g_4|o(g_4),d_2} f(o(g_4)))} / \frac{\hat{P}_{g_1|d_1}/(\overline{S}_{g_1|o(g_1),d_1} f(o(g_1))}{\hat{P}_{g_2|d_1}/(\overline{S}_{g_2|o(g_2),d_1} f(o(g_2)))}) = \frac{(\theta_{g_3} - \theta_{g_4}) - (\theta_{g_1} - \theta_{g_2})}{\sigma} \tag{12}$$

Thus, the adjusted log odds ratio identifies the expected gain in scaled joint surplus from swapping partners in any two job transitions. Note that differencing and conditioning, respectively, necessarily remove any information about the mean payoffs or welfare of worker types and position types. However, the identified set of surplus difference-in-differences $\Theta^{D-in-D} \equiv \{\frac{(\theta_g - \theta_{g'}) - (\theta_{g''} - \theta_{g'''})}{\sigma} \forall (g, g', g'', g''') : o(g) = o(g''), o(g') = o(g'''), d(g) = d(g'), d(g'') = d(g''')\}$ preserves the critical information about the relative efficiency of alternative matchings present in the observed group frequencies .

3.2 Counterfactual Simulations

In this subsection I show that identification of Θ^{D-in-D} is sufficient to generate the unique counterfactual aggregated assignment $P^{CF}(g)$ following arbitrary changes in the marginal distributions of worker origin types $f(o)$ and destination position types $h(d)$. If more than one matching is observed, then σ can be (roughly) estimated as well, permitting a proper welfare analysis of the

[19] d_2 could be (but need not be) the same destination type as d_1.

approximate mean utility and profit gain for each worker origin type and position destination type.

One can characterize the set of workers to be reallocated via their origin type distribution, $f^{CF}(o)$. The "CF" superscript indicates that this distribution could be counterfactual (e.g. capturing a proposed influx of refugees). Similarly, the set of counterfactual positions to be filled can be represented by $h^{CF}(d)$, and the prevailing matching technology can be denoted $\{\theta_g^{CF}\}$. The values $f^{CF}(o)$, $h^{CF}(d)$, and $\{\theta_g^{CF}\}$ are all inputs that are either observed or chosen by the researcher, who wishes to predict the equilibrium aggregate distribution of transitions, $P^{CF}(g)$.

To take a concrete example, suppose that a local development board has already forecasted the number and location of new manufacturing positions (and perhaps spillover-driven retail positions) that a plant opening would generate, and suppose the existing group mean surpluses $\{\theta_g^{CF}\}$ have been estimated. The board may wish to predict the extent to which the plant opening will increase job-related utility and the employment rate among existing local workers/job seekers in the chosen and surrounding neighborhoods (and perhaps the profits of local and less local firms).

I assume that the counterfactual assignment also satisfies the assumptions of Proposition A1 above. I also assume that the set of destination type averages of the shares of workers of each origin type who would be incumbents at the establishment, $\{\overline{S}_{g|o,d}^{CF}\}$, 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 $\overline{S}_{g|o(g),d(g)}^{CF} = \overline{S}_{g|o(g),d(g)}^{y'} \ \forall\ g$, which is observed. Then the counterfactual conditional choice probability $P^{CF}(g|d)$ can be expressed as (10) with $(\theta_g^{CF}, f^{CF}(o), h^{CF}(d), \overline{S}_{g|o(g),d(g)}^{CF}, C_o^{CF})$ replacing $(\theta_g, f(o), h(d), \overline{S}_{g|o(g),d(g)}, C_o)$. The worker type-specific mean exponentiated (and rescaled) utility values $\mathbf{C}_o^{CF} \equiv \{C_1^{CF} \dots C_O^{CF}\}$ are ex ante unknown equilibrium objects affected by the counterfactual changes reflected in $(\theta_g^{CF}, f^{CF}(o), h^{CF}(d))$. Thus, each counterfactual CCP must initially be treated as a function of the vector $\mathbf{C}_o^{CF}$.

GS and Decker et al. (2013) each show that a unique probability distribution over transition groups $P^{CF}(g)$ satisfies the aggregate analogues to the stability and feasibility conditions. However, these papers as well as CS assume when proving identification that one observes the numbers of unmatched partners (and thus the total number of agents) of each type on both sides of the market. While counts of unemployed workers by type can be accurately constructed, the LEHD database contains no information about unfilled vacancies.[20] Because each submatching of a stable matching must also be stable, observing only the subset of positions that match with workers does not threaten identification of the remaining elements of Θ^{D-in-D}; the estimated surplus relationships would not be reversed if data were augmented with 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 job assignment. Furthermore, unfilled positions might potentially choose to become filled when wages fall and previously filled positions might be left unfilled when wages rise in the wake of labor demand shocks. Ignoring this margin of adjustment might cause my incidence estimates to slightly overstate the

[20]Furthermore, constructing counts of vacancies for the destination types used in this paper from publicly available vacancy data is also not straightforward.

magnitude of welfare gains from positive shocks and losses from negative shocks. To rule out such scenarios, I assume that the costs of adjusting the number of positions at establishments (and thus changing the assignment of production tasks to workers) are large relative to the changes in the minimized cost of an efficiency unit of labor from the relatively small and localized shocks that I consider. In this case, the shocks will only cause establishments to adjust the worker composition at a pre-set number of positions at each location, so that $h^{CF}(d)$ may be considered fully exogenously determined.[21] Similarly, I assume that the second-best options for all workers who accept destination jobs consist of other positions that settle for other workers rather than positions that remain vacant, ensuring that unfilled vacancies ignored by $h^{CF}(d)$ also do not affect the division of surplus.

This perfect inelasticity assumption, while strong, simplifies the choice of variation used to identify of relative surplus values. One need not isolate labor supply shocks in order to identify extensive margin labor demand elasticities by type. Instead, surplus diff-in-diffs Θ^{D-in-D} (along with σ) are essentially determining equilibrium elasticities of substitution for each destination position type among different origin worker types. Elasticities of substitution are fully determined by *relative* prices, so they should be insensitive to the source of relative cost changes for different worker types: upward (downward) shifts in the number of local (distant) workers seeking positions and downward (upward) shifts in the number of local (distant) positions tending to prefer local (distant) workers are all valid sources of variation in relative prices of workers from different initial locations. So there is no inconsistency in using the full set of year-to-year job flows, implicitly driven by a mix of many small and large local supply and demand shocks, to recover Θ^{D-in-D}.

The perfect inelasticity assumption also dramatically simplifies the otherwise demanding computation of counterfactual equilibria. With an unknown number of unmatched partners on each side, GS show that one must solve $O + D$ non-linear equations that combine the feasibility and stability conditions for the mean equilibrium payoffs of all worker and firm types ($\{C_o^{CF}\}$ and $\{C_d^{CF}\}$). By contrast, when the "supply" of positions by destination type is assumed known, these values can be set equal in equilibrium to worker "demand" for such positions to create a set of D type-level market clearing conditions that determine $\{C_d^{CF}\}$.[22] Equivalently, if a dummy "position" type is added with mass equal to the share of workers who will end up unmatched, then the augmented demand (including "demand" from nonemployment) for each origin type o will equal the supply $f^{CF}(o)$, facilitating market clearing on the worker side.[23] Since relative payoffs among origin types fully determine the equilibrium assignment (so one can normalize $C_1^{CF} = 0$), and the origin-type distri-

[21] For larger shocks that make this assumption untenable, one could use existing wage elasticity and multiplier estimates to incorporate the endogenous response into $h^{CF}(d)$ and re-interpret it as a post-adjustment distribution.

[22] 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 to construct the stable assignment.

[23] These dummy nonemployment positions represent a computational mechanism for incorporating workers' payoffs from nonemployment, $\{\pi_{i0}^i\}$, akin to "balancing" an unbalanced assignment problem (Hillier and Lieberman (2010)). A formal proof of the equivalence of these sets of market clearing conditions is available from the author upon request.

bution $f^{CF}(*)$ must sum to one, we obtain the following $O-1$ market clearing conditions:

$$\sum_{d \in \mathcal{D}} h^{CF}(d)(\sum_{g:o(g)=2} P^{CF}(g|d, \mathbf{C^{CF}})) = f^{CF}(2)$$
$$\vdots$$
$$\sum_{d \in \mathcal{D}} h^{CF}(d)(\sum_{g:o(g)=O} P^{CF}(g|d, \mathbf{C^{CF}})) = f^{CF}(O) \tag{13}$$

Given a solution to (13), one can then construct the counterfactual transition probability for any transition group via $P^{CF}(g) = \sum_d h^{CF}(d) P^{CF}(g|d, \mathbf{C^{CF}})$. Since the solution also satisfies the stability and feasibility conditions, it must be the unique aggregate counterfactual stable assignment.

Because only $\min\{O, D\}$ equations must be solved, this approach provides considerable computational savings when the number of types is much larger on one side of the market. Below I present results that average over 500 counterfactual allocations featuring around 1,000 origin and 10,000 destination types that would be prohibitive to compute with unmatched agents on both sides.

3.3 Interpreting the Counterfactual Simulations

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

First, constructing the market-clearing conditions (13) requires a full set of group joint surpluses $\Theta^{2010} \equiv \{\theta_g^{2010} \, \forall \, g \in \mathcal{G}\}$, but the identification argument in section 3 suggests that only the set of diff-in-diffs $\Theta^{D-in-D,2010}$ is identified. In Appendix A2, I prove the following proposition:

Proposition 1:

Define the set $\Theta^{D-in-D} \equiv \{\frac{(\theta_g - \theta_{g'}) - (\theta_{g''} - \theta_{g'''})}{\sigma} \, \forall \, (g, g', g'', g''') : o(g) = o(g''), o(g') = o(g'''), d(g) = d(g'), d(g'') = d(g''')\}$. *Given knowledge of* Θ^{D-in-D}, *a set* $\tilde{\Theta} = \{\tilde{\theta}_g \, \forall \, g \in \mathcal{G}\}$ *can be constructed such that the unique group level assignment* $P^{CF}(g)$ *that satisfies the market-clearing conditions (13) using* $\theta_g^{CF} = \tilde{\theta}_g \, \forall \, g$ *and arbitrary marginal PMFs for origin and destination types* $f^{CF}(*)$ *and* $g^{CF}(*)$ *will also satisfy the corresponding market-clearing conditions using* $\theta_g^{CF} = \theta_g \, \forall \, g \in \mathcal{G}$ *and the same PMFs* $f^{CF}(*)$ *and* $g^{CF}(*)$. *Furthermore, denote by* $\mathbf{\tilde{C}^{CF}} \equiv \{\tilde{C}_1^{CF}, \ldots, \tilde{C}_O^{CF}\}$ *and* $\mathbf{C^{CF}} \equiv \{C_1^{CF}, \ldots, C_O^{CF}\}$ *the utility vectors that clear the market using* $\theta_g^{CF} = \tilde{\theta}_g$ *and using* $\theta_g^{CF} = \theta_g$, *respectively. Then* $\mathbf{\tilde{C}^{CF}}$ *will satisfy* $\tilde{C}_o^{CF} = C_o^{CF} e^{\frac{-\Delta_o}{\sigma}} \, \forall \, o \in \mathcal{O}$ *for some set of origin type-specific constants* $\{\Delta_o : o \in [1, O]\}$ *that is invariant to the choices of* $f^{CF}(*)$ *and* $g^{CF}(*)$.

Essentially, the proposition states that the identified set of surplus difference-in-differences Θ^{D-in-D} contains sufficient information to generate the unique counterfactual group-level assignment $P^{CF}(g)$ associated with the complete set of surpluses Θ. Furthermore, the utility premia $\tilde{\mathbf{C}}^{\mathbf{CF}}$ that clears the market using the artificially completed surpluses $\tilde{\Theta}$ will always differ from the "true" premia $\mathbf{C}^{\mathbf{CF}}$ that clear the counterfactual market under Θ by the same o-type-specific constants regardless of the compositions of supply $f^{CF}(o)$ and demand $h^{CF}(d)$ that define the counterfactual.

The existence of the "bias" terms $\{\Delta_o\}$ in Proposition 1 indicates that even relative levels of utility among origin types in counterfactual allocations (or observed allocations) are not identified.[24] However, because the Δ_o values are constant across counterfactuals featuring different $f^{CF}(o)$ and $h^{CF}(d)$ distributions, the relative *changes* $[(\ln(C_o^{CF1}) - \ln(C_o^{CF2})) - (\ln(C_{o'}^{CF1}) - \ln(C_{o'}^{CF2}))] \approx (\frac{(\overline{r}_o^{CF1} - \overline{r}_o^{CF2}) - (\overline{r}_{o'}^{CF1} - \overline{r}_{o'}^{CF2})}{\sigma})$ in mean rescaled utilities across origin types among two counterfactuals can be identified.[25] Below, pairs of counterfactuals include one that features a demand shock and an otherwise identical counterfactual that does not. In some cases one may reasonably impose a priori values for utility changes for one or more origin types, so that utility changes $\frac{\overline{r}_o^{CF1} - \overline{r}_o^{CF2}}{\sigma}$ for other types can be identified. Below I assume that the small, very local stimuli and natural disasters I consider generate zero utility change for the least affected (usually quite distant) worker type. Such restrictions allow each worker type's share of total welfare gains (or losses) from a labor demand (and/or supply) shock to be determined. The model's symmetry between workers and positions implies that mean changes in profits and shares of profit gains (or losses) by destination type also are identified. Thus, given data on even a single matching, the model can produce a reasonably complete analysis of job-related welfare incidence from labor supply and demand shocks.

Second, in addition to these normalizations, in order for the predicted allocation and welfare gains to accurately represent what would have happened had the simulated shocks occurred, one must also assume that the joint surpluses Θ^{CF} and the marginal type distributions $f^{CF}(*)$ and $h^{CF}(*)$ that serve as simulation inputs are exogenous to (i.e. unaffected by) the shock itself. Any reallocation and welfare changes are assumed to be driven exclusively by the changes in transfers across origin types required to eliminate the shock-induced imbalances between supply and demand.

As discussed above, exogeneity of $h^{CF}(*)$ imposes that the shock does not cause further changes in firms' location and size decisions. To highlight heterogeneity in the scope of labor markets by firm size, average pay, and industry, below I only consider simple "apples-to-apples" comparisons where each stimulus package involves adding a common number of jobs to a single destination type. However, one could incorporate agglomeration and congestion forces to customize the simulations by adding or subtracting positions for other types to the original "exogenous" shock in accordance with existing job multiplier estimates from the literature (e.g. Bartik and Sotherland (2019)).

There are also plausible mechanisms by which the joint surpluses Θ^{2010} might respond to the

[24]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 CS or GS, stems from the lack vacancy data.

[25]This approximation requires that the variation in utility values among workers assigned to the same origin type is limited, so that $\ln(C_o) \equiv \ln(\frac{1}{|o|}\sum_{i:o(i)=o} e^{\frac{r_i}{\sigma}}) \approx \ln(\frac{1}{|o|}\sum_{i:o(i)=o} e^{\frac{\overline{r}_o}{\sigma}}) = \frac{\overline{r}_o}{\sigma}$.

shock.[26] However, for reasonably small local shocks, the most obvious endogenous surplus changes are likely to be minuscule relative to the size of existing surplus variation in the worker types' relative productivities across firm types, relative valuations of firm and location amenities, and moving costs from alternative transitions, so that such exogeneity violations generate minimal bias. Note also that only changes in surplus diff-in-diffs Θ^{D-in-D} affect the counterfactual assignment, so that the components of endogenous changes to productivities, amenities, or continuation values among position types that are common to all workers do not affect the shock's incidence among workers.[27]

Another caveat relates to a shock's duration. I focus on forecasting reallocations and welfare changes that occur within one year of the shocks, and I assume that transitions to relocated and stimulus-generated positions create the same surplus as those to existing positions of the targeted destination type. Implicitly, this requires that the new positions have the same expected duration as other positions of their type. To evaluate a particularly temporary construction stimulus, for example, one might wish to estimate separate surplus parameters for temporary and permanent construction jobs.[28] As is, the model is designed to show that the incidence of even very local shocks may spread quite widely across space (and skill levels) even over a short period, despite strong tendencies consistent with large short-run mobility frictions of movers to take nearby jobs.

A final, important caveat relates to the absence of a housing market in the model (and residential choices in the data). Standard models of spatial equilibrium in urban economics (e.g. Roback (1982) or Kline and Moretti (2013)) emphasize that if housing supply is quite inelastic and workers are very mobile, a large share of the incidence of a positive place-based shock accrues to landholders via higher rents (which offset the utility gains to workers from any wage increases). However, this same argument suggests that for local workers who are also nearby renters, the simulated employment-related welfare gains are likely to place an upper bound on the full welfare gains these workers would experience.[29] Thus, local low-skill workers would be well justified in resisting local initiatives focused on bringing "good" jobs to town if they are likely to generate an employment-related incidence that is either geographically dispersed or concentrated among higher skilled workers.

[26]For example, the existence of a new establishment nearby might increase the demand for intermediate products produced by other local firms, raising the productivity of workers for such firms. Alternatively, if the new positions are permanent and search/recruiting/moving costs increase with distance, then jobs at nearby establishments might now have greater continuation value because future job searches will begin in a local area featuring greater labor demand.

[27]Such composition-induced surplus changes also will not affect the equilibrium origin type payoffs that capture the shock's incidence among groups of workers. From the proof of Proposition 1 in Appendix A2, we see that such surplus changes only change Δ_d^2, which does not enter into equilibrium mean payoffs for origin types $\{C_o^{CF}\}$. Instead, any increase in joint surplus for a destination position that is common to all workers will be fully reflected in its profit payoff, through higher revenue for the same costs or lower salaries that offset the change in worker continuation values. The profit gains among nearby position types will be understated, however. This partly motivates the focus on incidence among workers, for whom differential agglomeration effects among firms across shock compositions may be less important.

[28]Fully elucidating the differences in welfare effects between shocks of different expected durations requires a fully dynamic assignment model along the lines of Choo (2015) that specifies both worker expectations and the serial correlation in (now time dependent) idiosyncratic surplus components ϵ_{ikt}. would need to be specified. Similarly, the Markov-style model used here cannot accommodate optimal location sequences and return migration choices highlighted by Kennan and Walker (2011). See Weinstein (2018) for an evaluation of the dynamic consequences of a particular local shock.

[29]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.

Furthermore, house prices may not change much in places where housing supply is likely to be relatively elastic (such as rural areas or areas with weak zoning laws) or where commuting costs are low (so that adjustment to shocks occurs primarily via changing commuting patterns). In these cases, abstracting from the housing market may produce minimal bias in incidence forecasts. Indeed, since commuting costs from job transitions involving location changes are implicitly captured in the model as a component of the joint surplus θ_g, they will be reflected in incidence estimates.[30] While a complete welfare analysis requires incorporating housing and product markets, the goal of this paper is to highlight the heretofore underappreciated roles of differential geographic scopes of local labor markets for different types of workers and firms and the skill vs. spatial mismatch tradeoff in determining the incidence of alternative local labor demand interventions.

3.4 Identifying σ

While the share of welfare gains or losses for workers (or firms) can be identified without knowledge of σ, it is nonetheless a parameter of interest. Since payoffs are additive in worker earnings, knowledge of σ allows the estimated utility gains $\frac{\overline{r}_o^{CF1}-\overline{r}_o^{CF2}}{\sigma}$ and profit gains $\frac{\overline{q}_d^{CF1}-\overline{q}_d^{CF2}}{\sigma}$ to be re-scaled in dollar terms, making it easy to gauge whether the welfare changes from various labor market shocks are economically meaningful. Conditional on Θ, σ governs the elasticity of matching choices with respect to relative wages or required utility bids, so that it determines that magnitude of utility reallocations across skill types and locations following changes in labor demand composition.[31]

As Galichon et al. (2017) have noted, identification of σ requires combining information from multiple matchings, so I estimate σ using observed matchings from 2002-2003 and other years between 2003-2004 and 2010-2011.[32] Because the procedure (described fully in Appendix A3) requires additional strong assumptions, the estimate of σ is likely to be quite rough.[33] In practice, the estimates I obtain for σ are surprisingly consistent across years. I use the mean estimate of σ^y across all sample years, $\overline{\sigma} = 19,600$, to assign dollar values to all utility gains presented below.

As noted by GS, in the CS assignment model observed earnings in destination job matches

[30]Differential willingness to pay for locational amenities will be reflected in the relative propensities for different worker types to move to positions at particular locations, which are captured by the odds ratios used to identify Θ^{D-in-D}.

[31]Intuitively, when destination position type C disproportionately chooses type A workers over type B workers compared to position type D, it could be because $\theta_{AC} - \theta_{AD} \gg \theta_{BC} - \theta_{BD}$ and σ is substantial, or because $\theta_{AC} - \theta_{AD}$ is marginally larger than $\theta_{BC} - \theta_{BD}$ but σ is tiny. When the former is true, large changes in required utility bids are necessary to engender sufficient substitution across worker types to overcome strong comparative advantages from matching certain types of workers and positions. If the latter is true, small utility changes suffice to clear the market after a shock.

[32]The existence of a single σ parameter governing relative wage elasticities stems from assuming additive separability of θ_g and ϵ_{ik} and an i.i.d type 1 extreme value distribution for the vector of ϵ_{ik} values.

[33]Essentially, differences among origin types in their observed mean annual earnings changes between years $y-1$ and y are regressed on model-generated log differences in predicted scaled utility values $\ln(C_o^{CF,y}) - \ln(C_{o'}^{CF,y}) \approx (\overline{r}_o^{CF,y} - \overline{r}_{o'}^{CF,y})/\sigma^y$. These predicted values are constructed by computing counterfactual allocations in which origin and destination type distributions evolve as they actually did but joint surplus values are held fixed at their estimated 2002-2003 values. Under the assumptions that a) the 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 Θ, and b) mean utility gains for each origin type generally took the form of increases in annual earnings in the chosen year rather than increases in amenities or continuation values, the coefficient on $(\overline{r}_o^{CF,y} - \overline{r}_{o'}^{CF,y})/\sigma^y$ will approximately equal σ^y.

also can be used to decompose each group-level joint surplus value θ_g into the sum of worker and position subcomponent θ_g^l and θ_g^f (group-level analogues of π_{ik}^i and π_{ik}^k). However, in Appendix A4 I show that clean identification of θ_g^l and θ_g^f breaks down without the particular structure CS place on the unobserved match quality component ϵ_{ik} unless further strong assumptions are imposed. I do not pursue this approach because I have shown that this decomposition not needed to determine the incidence across worker and position types of alternative local labor demand shocks.

4 Data

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 records collected for unemployment insurance purposes containing quarterly job earnings and unique worker and establishment IDs for a near universe of jobs in the state.[34] The worker and establishment IDs are then linked across states, and the data are augmented with firm and establishment characteristics (notably establishment locations and industry codes) from an extract of the ES-202/QCEW report and worker demographics from the Social Security Administration (including age, race and sex but not occupation nor education for most workers).[35]

4.1 Sample Selection

My sample consists of all LEHD records from the 19 U.S. states that opted to provide data to my FSRDC project. By agreement with the Census Disclosure Avoidance Review staff, 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. While some states begin reporting data as early as 1990, others begin as late as 2003. While the estimation of σ and the model validation exercise described in Section 6.3 use all the data after 1997 and after 2003 respectively, the simulations of counterfactual shocks are based on surplus parameters estimated from 2010-2011 data. Preliminary work suggested that the shock incidence forecasts were quite insensitive to the pair of years chosen.[36]

To transform person-year observations to year-to-year transition/retention observations, I first define each individual's primary job in each year as the one featuring the highest earnings that exists for at least one full quarter, then aggregate earnings from the primary job across all quarters within the year, and then append the primary job from the following year to the current observation.[37] A worker who does not report earnings above $2,000 at any job in any full-quarter in 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

[34] 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.

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

[36] This was true despite the decreasing job-to-job mobility over this time period documented by Hyatt et al. (2016).

[37] A job is observed in a full quarter if it features positive earnings in the preceding and following quarter as well.

observed as employed in one of the sample states in any of the years for which data was provided.

I limit each individual's presence in the sample to begin and end with his/her first and last years of observed employment. Thus, the sample initially only includes spells of nonemployment that are bookended by 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. I also restrict the sample to person-years featuring individuals with ages between 20 and 70.[38]

Only observing 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 transition groups defined by this combination of characteristics are coarser than those in the model, I use the (sometimes spurious) U-to-E, E-to-U, and U-to-U transitions in the LEHD only to distribute the ACS group counts across the finer groups in the model. I supplement the ACS data with BLS data on national unemployment rates by age group to make the scale of the labor force consistent with standard measures. Appendix A5 provides further detail about this imputation procedure.[39]

The second problem is that excluding job-to-job transitions into or from states outside of the 19 state sample causes the counterfactual simulations to overstate the geographic concentration of demand shock incidence, since workers from these states are excluded from competing for in-sample positions. Because accurate forecasts of shock incidence are likely to be particularly sensitive to observing flows of workers to and from states adjacent to the one receiving the shock, in the simulations below I only choose target census tracts from a subset of 10 states in the west/southwest/great plains area where coverage is nearly complete, so that almost all adjacent states are observed (though I always use the full 19 state sample to represent the "national labor market" to minimize the remaining bias).[40] Despite overstating the within-state share of shock incidence, below I consistently find that a large share of welfare incidence accrues to initially out-of-state workers.[41]

[38]This limits the influence of "nonemployment" spells consisting of full-time education or retirement followed by part-time work, so that parameters related to unemployment are identified primarily from prime-aged workers who were unemployed or temporarily out of the labor force. In addition, for disclosure avoidance reasons, the results in this draft are based on a 50% random sample of all transition-level observations from the sample just defined.

[39]This imputation procedure also solves an additional problem: by only including spells of nonemployment that are bookended observed employment, my LEHD sample in isolation will 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 yet found jobs that end their unemployment spells. However, under the imputation procedure, the ACS and BLS data rather than biased LEHD counts set the scale of U-to-E, E-to-U, and U-to-U transitions (albeit at a slightly aggregated level).

[40]Using ACS 1-year residential mobility data and weighting states by their census tract count, I estimate that for the 10 states chosen to supply target tracts, about 47% of year-to-year worker inflows from other states and about 92% of total job-to-job changes ending in one of these 10 states (including within-state flows) originate in one of 19 in-sample states.

[41]An alternative approach would be to use the ACS to construct counts for a set of course transition groups featuring

4.2 Assigning Job Matches to Types and Job Transitions to Groups

For each pair of years $(y-1, y)$ I assign each transition/retention observation to an origin type $o^{y-1}(i)$, a destination type $d^y(k)$, and a transition group $g^{y-1,y}(i,k)$ (time superscripts will henceforth be dropped except where necessary). Specifically, workers are assigned to origin types based on the combination of their year $y-1$ primary establishments' locations (discussed further in Section 5.2 below) and the earnings quartile associated with their earnings from this establishment.[42] For workers who were not employed in $y-1$, the location of their most recent establishment is used (or, for new entrants, the location is imputed using ACS/LEHD data) and the earnings quartile is replaced by one of two unemployed categories, differentiated by age (< 25 or ≥ 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. Workers' year y primary positions are assigned to destination types based on the combination of the relevant establishment $m(k)$'s geographic location, industry supersector, and size and average worker earnings quartiles based on establishment-level employment and average pay distributions. These characteristics were chosen because they are consistently observable and likely to be important determinants of productivity complementarities, recruiting, search and moving costs, and the other components of the job match surpluses heterogeneous positions will create with heterogeneous workers. The transition/retention (i,k) is assigned to a group $g(i,k) \equiv g(o(i), d(k), z(i, j(i), k))$ based on the origin worker type $o(i)$, the destination position type $d(k)$, and an indicator $1(m(j(i)) = m(k))$ for whether the transition represents a retention or a job change.

4.3 Summary Statistics

Figure 1a (Table A1, Col. 1) displays a histogram of the distance between the locations of origin and destination establishments for workers who changed primary jobs $(m(j) \neq m(k))$ between 2010 and 2011. 3.5% of job switchers took new jobs within the same census tract, while another 7.6%, 7.2%, and 13.8% percent moved to jobs one, two, or 3+ tracts away within the same PUMA. 60% found jobs in another PUMA within the same state, while 8% changed states. The sizable share of workers accepting new jobs very near 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 workers' local labor markets may still hold outsized importance for their job-related welfare.

Table 2 Panel A, Row 1 shows that about 21.3 percent of year-to-year pairs of primary jobs in the sample involve moves to new jobs (either E-to-E transitions or U-to-E transitions). A full 70.3% of workers keep the same primary job from one year to the next, while 4.2% and 4.3% of year-to-year

out-of-sample states. I hope to pursue this approach in a future version.

[42]Earnings quartile cutoffs are defined relative to the distribution of primary job annual earnings among workers in the observation's state-year combination, and are based on prorating earnings from full quarters only to ensure that the quartile captures a worker's skill rather than the share of the year he/she worked. A worker's location must be imputed for multi-establishment firms. However, the Census Bureau's unit-to-worker imputation procedure assigns establishments with probabilities that depend on the distance between that establishment and the worker's residence, so any mistakes will likely misattribute the worker's job to another nearby establishment, limiting scope for significant measurement error.

pairs involve U-to-U and E-to-U transitions, respectively. Collectively, the 2010-2011 estimation sample for the set $\{\Theta^{D-in-D}\}$ features 23 million transitions and retentions.

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 need to consider these two groups of unemployed workers separately. We also see that employed workers in the lowest earnings quartile in 2010 were far less likely than higher paid workers to stay at their job (66.7%) and far more likely to transition to unemployment (9.7%) or another job (23.6%); 87.3% of highest quartile workers in 2010 were retained, while only 10.2% transitioned to new jobs. However, conditional on changing jobs, the highest earners were most likely to leave their original PUMA and to change states, suggesting that the geographic scope of labor markets may differ across skill categories. These differences motivate the use of nonemployment status/earnings quartiles as characteristics defining origin worker types.

Panel B of Table 2 shows that establishments in the highest average pay quartile are considerably more likely to retain their workers, but are also slightly more likely to hire more distant workers when filling a vacancy: 9.6% of their new workers were previously working out of state and 29.2% were working in the same PUMA, compared to 7.5% and 32.3% for the lowest paying establishments. The largest quartile of establishments (based on employment) are also more likely to retain their workers, but are the least likely to hire from out of state (4.3%). The smallest establishments hired a whopping 27.2% of new workers from out of state and only 28.2% from the same PUMA, suggesting, perhaps surprisingly, that the smallest establishments seem to operate in the most geographically 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 hold all but one establishment characteristic fixed 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 two kinds of local labor demand shocks, stimulus packages and natural disasters.[43] Each stimulus shock consists of 500 jobs that are added to the destination-year stock of positions 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 stimulus packages featuring different kinds of new establishments represented by combinations of the non-location establishment attributes that define a destination type: quartiles of establishment size and average

[43] I also experimented with plant relocations that moved jobs to a new location from a distant state. However, these shocks had virtually identical employment and welfare incidence to stimulus analogues featuring the same shock composition across all origin types within the receiving state. A table of relocation simulation results is available upon request.

pay along with industry supersector. Table 1 details the 35 shock compositions, which were chosen to highlight the heterogeneity in incidence across different industry/size/avg. pay cells. The final three packages add a requirement that the new positions may only be filled by workers from the surrounding PUMA, reflecting the kinds of stipulations included in some economic development contracts between cities and incoming firms.[44] Comparing these "restricted" specifications to their unrestricted counterparts illustrates the value of these provisions to cities or states.

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 lost jobs added as nonemployment "positions". These simulations reveal whether the skill and spatial incidence of negative shocks is symmetric to positive shocks, and also illustrate the degree to which higher skilled workers initially working in the targeted tract disproportionately obtain any local jobs that remain. These disaster simulations also demonstrate how the two-sided matching model could be customized to assess the labor market impact of any particular disaster scenario, including hurricanes that hit several contiguous tracts simultaneously (and perhaps with differential force).

5.2 Collapsing the Type Space for Distant Geographic Areas

Since transition groups g are defined by several other worker and establishment characteristics in addition to origin and destination locations, treating all 28,000 census tracts in the 19 state sample as separate locations would generate trillions of transition groups. Given the particular interest in the incidence of alternative demand shocks across locations relatively close to the shock, I combine initial types (and thus groups) that share the same worker and establishment characteristics and are geographically proximate to each other but far from the shock. Specifically, outside a 5-tract circle surrounding the targeted tract a type's location is defined by a PUMA rather than a tract, and outside the targeted state a type's locations is defined by a state rather than a PUMA.

Coarsening the type space for distant geographic locations dramatically decreases the overall number of groups and the sparsity of the empirical group distribution $\hat{P}(g)$. While many job-to-job transitions are between nearby tracts, very few transitions occur between most distant pairs of tracts, so that relative surplus parameters for transition groups featuring tracts in different states would be weakly identified without such coarsening. Note that this approach still incorporates all the available transitions and all 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 combining types, there are relatively few observed transitions per group g, particularly for groups local to the shock. Thus, following Hotz and Miller (1993) and Arcidiacono and Miller (2011), I smooth $\hat{P}(g)$ prior to estimation by replacing each element's value with a kernel-density weighted average of $\hat{P}(g)$ among groups featuring "similar" worker and position characteristics. Because excessive smoothing across other groups erodes the signal in the data about the degree of heterogeneity in the relative joint surpluses from job matches featuring different worker

[44]For example, Empowerment Zones only subsidize wages for employees that are local residents (Busso et al. (2013)).

and position characteristics and locations, particular attention was devoted to the smoothing strategy.

Appendix A6 details the customized smoothing procedure. It is based on the intuition that the destination establishment's location is likely to be critical in determining the origin locations from which worker transitions create the most surplus (i.e. least moving/search cost), while non-location characteristics (size, average pay, and industry) matter more than location for determining the worker skill category (proxied by initial earnings quartile) that generates the most surplus.

Note that the type aggregation and smoothing procedures imply that type and group spaces change whenever a new target tract is selected. Furthermore, FSRDC disclosure restrictions prevent the release of results that are specific to a particular substate geographic location. Thus, I only report averages of incidence measures across 500 simulations for each shock type, where each simulation targets a different randomly chosen tract from the 10 state southwest/west/great plains subsample.[45]

After the simulations have been run, transition groups are again redefined in order to average simulation results across alternative targeted 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 surrounding the shock.[46] Note that during the simulations the spatial links between adjacent and nearby tracts are left entirely unrestricted. Thus, to this point no a priori assumption about the role of distance has been imposed by the model beyond the initial aggregation of distant tracts to PUMAs and states described above.

5.3 Inference

Given that I observe the universe and not a sample of job transitions within the available states, it is unclear how to define the relevant population for the purposes of inference. Furthermore, since I estimate nearly a million surplus parameters $\theta_g \in \Theta$, and each counterfactual incidence statistic depends on the full set Θ, any confidence intervals should provide information about the precision of the incidence forecasts as opposed to specific parameters. Rather than properly characterizing sampling error in isolation, I rely on the model validation results presented in section 6.3 to assess the combined contribution of sampling error and misspecification to out-of-sample forecast accuracy.[47]

[45] A census tract is only eligible to be a target tract in the simulations if it features at least 100 jobs, so that the parameters governing local firms' and workers' choices are well-identified. The same set of 500 randomly chosen target tracts is used for each shock specification to facilitate fair comparisons among alternative stimulus packages and disasters.

[46] When defining types and presenting results I mostly focus on distance bins defined by tract, PUMA, and state pathlengths, since the number of potential workers contained within circles defined by the same pathlength is likely to be more consistent across urban and rural areas featuring very different densities than circles with a miles-based radius.

[47] The one exception is that in the first few results tables, I provide standard errors that reflect the sampling error stemming from averaging over only one possible draw of 500 census tracts, while the population parameter of interest is the average among all tracts in the 19 state sample (around 28,000). These standard errors are tiny, suggesting little value to running more than 500 simulations per specification. As a result, subsequent tables do not report standard errors.

6 Results

6.1 Stimulus Packages

6.1.1 Incidence by Distance to Focal Tract

Before comparing stimulus packages of 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 changes in assignments across all 32 stimuli simulated, effectively integrating over the joint distribution of establishment industries, sizes, and average pay levels. While I focus attention on figures containing graphical representations of the results, most figures have an accompanying Appendix table (listed in parentheses) that contains the plotted values.

Figure 2a (Table A1, Col. 3) displays the mean probability of receiving one of the 500 new stimulus jobs among randomly chosen individuals initially working or seeking employment at different distances from the focal tract. The figure highlights a sense in which U.S. labor markets are still quite local: the probability of obtaining a stimulus jobs for a worker initially within the target tract (.015) is three times higher than for a worker in an adjacent tract (.005), over 7 times higher than for a worker 2 tracts away (.002), and 15 times higher than for a worker initially 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 a stimulus job 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 or 2 or more states away, respectively.

However, the vast differences in $P(\text{new job }|\text{distance from target})$ present a misleading guide to the overall geographic incidence of new jobs. This is because the target tract initially contains an extremely small fraction of the population at risk of obtaining these jobs. Figure 1b (Table A1, col. 2) shows the share of workers in the 19 state simulation samples that are working/seeking jobs in each distance bin from the targeted tract before the stimulus. Only 0.0045% of the workforce is originally in the target tract. As expected, the shares quickly get larger for distance bins defined by concentric rings with larger diameters: 0.026%, 0.059%, 0.3% of the workforce is initially 1, 2, or 3+ tracts away from the target tract within the same PUMA, while 0.4%, 1.1% and 16.2% is 1, 2, or 3+ PUMAs away, respectively, and 13.5% and 68.4% is 1-2 states and 3+ states away.

Consequently, one obtains a very different impression of incidence by swapping the terms in the conditional probability and calculating the share of stimulus jobs obtained by workers initially working or nonemployed in each of the distance bins, $P(\text{distance from target }|\text{new job})$. Figure 2b (Table A1, Col. 4) displays the mean share of new jobs by distance bin across the 32 simulated stimulus packages. 3.1% of new jobs go to workers from the target tract, another 16.2% go to other workers in the PUMA, 60% are obtained by workers in different PUMAs within the state, and 10.6% go to out of state workers. So a very large share of the new jobs are likely to be taken by workers far from the local jurisdiction that receives the stimulus (and likely lobbies for its local placement).

One might 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 past plant openings. As emphasized in the introduction, the probabilities of obtaining the particular new jobs created by the stimulus package may not be very informative about the true incidence of the shock. This is because many workers who take the new jobs would have obtained other similar jobs in the absence of the stimulus, and nearby workers may now obtain these jobs, and so on, creating ripple effects through vacancy chains that determine the true employment and welfare incidence. This is where a flexible equilibrium model provides additional insight.

Figure 3a (Table A1, Col. 5) is analogous to Figure 2a (Table A1, Col. 3), except that instead of the probability of obtaining a stimulus job, it captures the change in the probability of any employment (i.e. the change in the employment rate) due to the stimulus, relative to a no-stimulus counterfactual, among those initially working/nonemployed at different distances from the target tract. The change in employment probability is still quite locally concentrated, but less so than the probability of obtaining a stimulus job. Workers from 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 changes in employment rates 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.

Figure 3b (Table A1, Col. 6), the employment rate analogue to Figure 2b (Table A1, Col. 4), displays the share of the aggregate 500 job increase in national employment that accrues to workers initially in each distance bin. Only 0.5% of the net employment change redounds to workers from the target tract, with 7.6% of employment gains going to workers in other tracts within the PUMA, 57.4% to workers in other PUMAs within the target state, and 34.5% to workers from out of state.

The simulation procedure also generates counterfactual changes in mean job-related utility for each origin type following the various stimuli.[48] Since only relative changes are identified, the estimated utility impact is normalized to 0 for the origin type estimated to experience the smallest impact (which varies by the stimulus composition, but is generally initially low paid workers in some distant state), so that all presented utility changes are relative to this origin type.

Figure 4a (Table A1, Col. 7) provides the average utility impact (scaled in annual earnings equivalents) by distance bin from the target tract for the "typical" stimulus package. Workers from the focal tract receive an estimated $1,045 increase (in 2010 dollars) in money metric utility from the typical stimulus package (relative to the least affected worker type), while workers initially 1, 2, and 3 or more tracts away receive expected utility gains of $395, $278, and $164 respectively. Workers initially 1, 2, and 3+ PUMAs away within the state receive the utility equivalent of $164,

[48]Recall, though, that scaling utility premia in dollars of annual earnings requires estimating σ. As discussed in section 3.4 and Appendix A3, the assumptions that underlie the estimate of σ are stronger than for the relative joint surplus values (they are more like approximations), so estimated dollar values of predicted welfare gains should be treated cautiously.

$143, and $109 in annual earnings gains, while workers one and 2+ states away receive average gains of $89 and $85. Figure 4b (Table A1, Col. 8) plots the share of total utility gains (relative to the normalized type) that accrue to workers in each distance bin. Only 0.5% of total worker welfare gains accrue to workers from the focal tract, with another 6.5% accruing to other workers originally within the PUMA of the focal tract. 51.5% of the gains accrue to workers outside the targeted PUMA but within the same state, while 41.6% go to workers from 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 A2 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 tract, PUMA, or state pathlength. The story is the same: only 10.3% of employment gains and only 11.6% of welfare gains accrue to workers 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 more than 250 miles away.

6.1.2 Heterogeneity in Incidence Across Establishment and Worker Characteristics

Another feature of the model is the ability to capture heterogeneity in incidence across workers in different skill classes, as proxied by initial employment status and earnings quartile. Figure 5a (Table A3, Col. 1) captures the share of the 500 job net employment gain enjoyed by workers whose initial earnings fall in each national quartile, as well as workers who were unemployed in the prior year. 41.1% of employment gains accrue to the initially unemployed (8.9% to workers younger than 25, 32.2% to older workers), while 26.1%, 15.4%, 9.6%, and 7.6% go 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 they were less likely to become unemployed in the absence of the shock. Figure 5b (Table A3, Col. 6) displays the share of total worker welfare gains enjoyed by unemployed workers and workers in each initial earnings quartile for the typical shock. 10.3% of utility gains go to initially unemployed workers, while the share accruing to each earnings quartile increases with initial earnings: 20.1%, 21.4%, 22.3%, and 26.0% for quartiles 1-4. Existing high paid workers seem to receive a disproportionate share of the welfare gains from a typical shock.

One can also consider the degree to which the geographic scope of labor markets depends on skill level. Figure 6a (Table A4) examines employment rate changes among workers whose initial jobs (or unemployment) place them in given earnings quartile/distance bin combinations. Older unemployed workers who most recently worked in the focal tract enjoy a sizable 1.6 percentage point decrease in their unemployment rate, while the decrease is only 0.4% and 0.3% for workers most recently employed one or two tracts away, indicating that employment gains for existing unemployed workers are quite local. That said, employment gains decline with distance in a somewhat similar pattern for all initial earnings quartiles. Existing older unemployed workers in the target tract enjoy 0.2% of the total employment gains despite constituting 0.0003% of the labor force.

Figure 6b (Table A5) displays the welfare analogue to Figure 6a (Table A4). Older unemployed workers from the targeted census tract enjoy a utility change equivalent to $1165 in 2010 earnings from a typical shock, while their younger counterparts realize 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 earnings quartile to $1242 for the 4th. Welfare gains decrease more quickly with distance for higher income groups, however, creating rapid convergence in gains across earnings quartiles with distance from the focal tract.

Figure 7a (Table A6, Col 2-9) shows the utility gains only for workers initially in the target tract 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 older unemployed workers equivalent to $1,374 and $1,413 in annual earnings, compared to $952 and $1000 for stimuli featuring new positions in the information or state/local government sectors.[49] Workers in the highest 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. The rankings of utility gains for local workers across industries 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 the lowest for initially high paid workers.

Figure 7b (Table A7) shows the expected utility gains for workers in the focal tract by establishment size and 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 unemployed workers. Stimuli featuring positions at small, high paying firms generate the least payoff for initially unemployed workers ($413 and $1015 for young 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 create larger gains for previously low-paid but employed workers while larger firms generate more gains for the initially unemployed.

In addition, the substantial heterogeneity in local skill incidence across industries and size/pay quartile combinations misses further heterogeneity at the three-dimensional supersector/size/pay cell level. Figure 8 (table available upon request) plots welfare gains by skill proxy among workers from the focal tract for all 32 stimulus compositions. The range of predicted gains is huge. Welfare gains for young 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 primarily concerned with very local incidence, these rep-

[49] "Other Services" includes repair and maintenance, personal and laundry services, and religious/civic organizations.

resent massive differences in the scale and skill intensity of utility incidence that would be obscured by an analysis that either ignored skill heterogeneity or used coarser geography.

On the other hand, city-level and particularly state- and federal-level policymakers may safely ignore incidence heterogeneity, since shocks featuring different establishment industry, size, and average pay composition exhibit increasingly similar spatial and skill incidence as one focuses on locations more than two or three tracts from the stimulus site. Aggregating across skill types and size/average pay combinations, Table A8, Col. 2-9 and Table A9, Col. 2-9 show that the change in employment rate and the share of total employment gains experienced by workers at different distances varies shockingly little across stimuli featuring different supersectors, with all industries displaying within-PUMA shares of employment gains between 7.4% (other services) and 8.8% (retail/wholesale trade) and within-state shares between 65.1% and 66.5%. Similarly, when averaging over locations and focusing on skill incidence, all supersectors feature shares of utility (employment) gains accruing to each skill category within 1% (4%) of the overall average across all stimuli (table available upon request). Analogous results for utility (also available upon request) display a similar uniformity among industries in spatial and skill incidence away from the shock.

Changing the establishment size/pay composition of stimuli also produces limited heterogeneity in geographic, and more surprisingly, skill incidence beyond tracts close to the site of the shock (Tables A10 and A11). 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 A3). By contrast, initially unemployed workers are predicted to take $\sim$ 24% of stimulus jobs at high paying firms versus $\sim$ 30% for low paying firms, and the initially high paid workers take $\sim$ 34% versus 7% of jobs from high versus low paying firms, suggesting that the skill incidence of the particular stimulus jobs understates the degree to which employment gains from labor demand shocks featuring a bias toward high skilled workers "trickle down" to unemployed workers.

Besides the establishment characteristics that constitute the stimulus, another source of heterogeneity in geographic incidence stems from the choice of focal tract. Figure 9 (Tables A12 and A13) provides the mean employment and welfare incidence among the least dense (most rural) and most dense (most urban) 100 tracts (generally containing $<$ 500 and $>$8000 residents, respectively) of the 500 receiving simulated shocks, as well as among the 100 tracts with smallest and largest pre-shock employment counts (generally $<$250 and $>$8,000 workers employed in the tract, respectively).

Both employment and welfare gains are more geographically concentrated for tracts with lower population density. The average share of employment gains enjoyed by workers initially within the targeted PUMA is 9.5% among the 100 most rural tracts versus 5.6% for the 100 most urban tracts (and 7.1% among all 500 tracts simulated). The measures of welfare incidence tell a similar story, with the expected 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 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 few initial workers relative to those featuring many (e.g. $1790 vs. $434 for those from the focal tract). Interestingly, despite their disproportionate gains, there are so few existing workers in the 100 smallest tracts relative to the largest 100 tracts that the latter set features a greater focal tract share of total welfare and employment gains (0.3% versus 0.7% for both outcomes).

Furthermore, focusing on rural/urban contrasts masks additional unexplained heterogeneity in incidence among the 500 randomly chosen target tracts. For each shock specification, the within-PUMA shares of employment gains range from below 5% to above 15% and the within-state shares range from below 35% to above 80%, though these ranges partly reflect sampling error given relatively few observed transitions per group g. Shares of welfare gains display even greater variation, with within-state shares (partly driven by state size) ranging from below 10% to above 90%.

Finally, Figure 10 (Table A14) illustrates the impact on stimulus incidence of requiring incoming establishments to fill their positions using only workers from the same PUMA as the tract receiving the stimulus. The unemployment rate falls by 1.4% instead of 0.3% for workers in the target tract when hiring is restricted to occur within PUMA, while the decrease in unemployment rate 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 existing local PUMA workers.

Restricting hiring also increases the expected money metric utility gains by seven fold ($995 to $6938) for workers from 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, local development initiatives such as empowerment zones that add stipulations restricting hiring or wage subsidies to only the local workers seem likely to generate a much more locally concentrated labor market incidence. 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.

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 skill categories, Figure 11a (Table A15, col. 1-4) shows that workers from the focal tract experience unemployment rate increases after one year of 2.9%, 7.3%, and 19.2% from the 25%, 50%, and 100% disasters, respectively. The share of new unemployment borne by workers from the local tract increases from 12.6% to 20.2% when 100% instead of 25% of local jobs disappear, suggesting that the employment incidence becomes increasingly spatially concentrated the more intense the local disaster (even when the disaster is still contained within the same tract) (Table A15, col. 5-8). As with the stimulus shocks, in one sense these results suggest that labor markets are very local: the change in unemployment rate among local workers is over 4,000 times larger than it would be in a frictionless, homogeneous world in which the expected share of lost jobs borne by local workers would be their share of the total workforce: 0.0045%.

Note that the local share of total employment change for each simulated disaster is far larger than for any stimulus package. Since most local workers would have been working (somewhere) in the absence of the shock (or are the long-term unemployed who produce little joint surplus from employment), there was an effective limit to how local the employment incidence of stimulus packages could be, forcing most employment gains to accrue in distant locations. Analogously, since unemployment rates are well below 20% in most locations before the simulated shock, very few would have become unemployed without the disaster, creating scope for large local losses. Since most positions retain their existing workers (appearing in the model as large joint surplus values for retentions), local workers have difficulty finding jobs. Thus, the model estimates reveal a natural asymmetry in spatial employment incidence between positive and negative local demand shocks.

Figure 11c (Table A15, col. 9-12) shows the average earnings-equivalent utility losses by distance bin for each disaster intensity. Expected utility 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 type). Because losing 100% of local jobs is fairly far outside the support of tract-level employment changes observed in the data, and the model assumes away employment expansions by establishments responding to cheaper local labor, the scale of these welfare losses should be considered somewhat skeptically. The losses fall dramatically to \$-235, \$-361, and \$-543 for workers from an adjacent tract, and then decrease slowly in magnitude to $\approx$ \$-115, \$-135, and \$-135 for those from outside the targeted state. However, because within-tract workers are such a small share of the work force, they still only account for 10%, 9.7%, and 8.3% of national welfare losses for the three disaster intensities, respectively (Figure 11d, Table A15, col. 13-16). While this local share is 20 times larger than its stimulus analogues, it is nonetheless fairly small. As before, over 80% of the utility 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 a broadly shared overall incidence.

Figure 12a (Table A16, col. 1-3) displays the share of all employment losses experienced by each skill category. For disasters featuring 25% local job loss, nearly 35% of employment losses are borne by those already unemployed, with the share falling monotonically from 27.8% to 8.9% for the 1st to the 4th initial earnings quartile. Thus, high skilled workers can insulate themselves from employment losses by taking jobs from less skilled workers, creating a cascade of sorts. However, for more severe disasters, the burden of employment loss is more equally shared, with only 31% accounted for by the initially unemployed, and 11% accounted for by the initially highest paid.

However, Figure 13a (Table A17, Col. 1-6), which examines employment incidence by distance and skill category jointly, paints a richer picture. Among those from the focal tract, the increase in the unemployment rate from the least severe (25%) disaster is actually larger for employed workers than initially unemployed workers: younger (older) unemployed workers experience a 0.2 (0.6) percentage point increase in end-of-year unemployment, while workers at initial earnings quartiles 1-4 experience unemployment rate increases of 4.4, 3.6, 2.8 and 2.2 percentage points, respectively. This is because the unemployed workers had the least to lose; they were fairly likely to be unemployed again without a disaster. However, among workers a tract away or further, the employment losses are greatest among the existing unemployed. This pattern becomes even more pronounced for the most severe (100%) disaster: initially unemployed older (younger) local workers experience 0.6 (1.6) percentage point increases in unemployment rate, as compared with 21.8 and 19.4 percentage point increases for local 1st and 4th earnings quartile workers (Figure 13b, Table A17, Col. 7-12).

When averaging across distance categories, welfare losses (Figure 12b, Table A16, col. 4-6) are more concentrated among the initially employed and particularly the higher paid, and are insensitive to disaster severity: 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 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% (100%) job loss, workers 1 state away experience utility losses of $-118 ($-136 to $-139) relative to the least affected worker type, regardless of initial skill (Figure 14a, Table A18). Differences in welfare losses are similarly small within all distance bins except the focal tract. Both younger and older unemployed local workers are predicted to lose the equivalent of around $300 in utility in the 25% disaster, while focal tract workers in earnings quartiles 1-4 are predicted to lose $-4,755, $-6,043, $-7,022, and $-7,531, respectively. For the 100% disasters, 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 job-related utility.

Finally, while quantifying the employment and utility incidence of disasters is important for allocating relief funds, policymakers and local communities also care about flows of migrants away from disaster sites. Thus, Figure 15a (Table A19, col. 1) displays, for each disaster intensity, the change in the probability of being employed at establishments in each distance bin for workers initially in the tract hit by the natural disaster. For the mildest disaster, the decrease in within-tract employment for workers from the focal tract is only 10.6%, despite a 25% overall decrease in local positions. This is in part because existing local workers retain a greater share of remaining jobs, but also because 29.5% of these workers would have taken jobs in other tracts even in the absence of the shock: when 100% of local jobs are eliminated, the decrease in the share of local workers staying in the tract is 70.5%. Relative to a counterfactual without a disaster, an additional 2.0%, 4.8%, and 12.2% of the workers initially in the focal tract end up employed in other tracts within the original PUMA in the 25%, 50%, and 100% disaster scenarios. Locations outside the PUMA but in the state take on an additional 5.2%, 13.1%, and 34.8% of those from the focal tract in the three scenarios, while out-of-state locations take on an additional 0.6%, 1.6%, and 4.3%, respectively. An additional 2.9%, 7.3%, and 19.2% become unemployed. Thus, while a relatively small share of workers find employment in nearby neighborhoods, this share increases quickly with disaster severity.

Figure 15c (Table A19) displays separate distributions of destination job locations for target tract workers in each skill category. Even in the most severe disaster, only an additional 1.8% (0.8%) of younger (older) initially unemployed workers take jobs away from the focal tract, relative to the disasterless counterfactual. The few that would have obtained local jobs remain unemployed instead. By contrast, since most high paid workers are retained or continue to work locally in the absence of the shock (only 12.4% would have left the focal tract), the 100% disaster engenders a very large mobility response for such workers: an additional 13.2% switch tracts within the PUMA (relative to the counterfactual), and an additional 47.8% and 7.1% switch PUMAs within state and switch states, respectively. 1st quartile workers were more likely to switch jobs anyway, so their corresponding increases are only 11.9%, 30.9%, and 2.6%, respectively. Relative to low paid workers, the mobility response for high paid workers is disproportionately muted for lesser disasters, though, because they are better at capturing the remaining local jobs (Figure 15b, Table A19).

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 job transitions driven by small firm expansions/contractions and worker retirements and preference or skill changes over the life cycle that generate considerable offsetting churn in the U.S. labor market. Thus, one might reasonably wonder whether parameters governing ordinary worker flows are capable of capturing the response to sizable, locally focused positive or negative shocks. To address this concern, I perform a model validation exercise in which surplus parameters estimated on pre-shock ordinary worker flows were used to forecast the reallocation of workers after actual local economic shocks observed in the LEHD sample. I evaluate model fit using the index of dissimilarity between the predicted and actual transition group distributions $P(g)$, and average this index across 514 shocks defined by tract-year combinations featuring large positive or negative changes in employment (at least 100 workers and at least 10% of the pre-shock tract employment level), without offsetting contemporaneous shocks to other tracts in the PUMA or shocks to the same tract in other years. Appendix A7 describes the exercise in detail, while Appendix Table A20 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 much more narrowly defined locations within the target PUMA (census tracts rather than entire PUMAs or states), the model would need 36.5% of transitions originating in the target PUMA to be reallocated to other groups g to perfectly match the true within-PUMA distribution. However, a large share of "incorrect" predictions involve either slight differences in destination tract within the same distance bin or slightly mismatched combinations of establishment size/avg. pay/supersector categories. When the group space is collapsed post-simulation so that origin and destination locations are defined by 42 distance bins from the target tract and non-location establishment characteristics are excluded (though worker skill categories remain), the share of transitions that must be reallocated across collapsed groups to match the data falls to 0.9% nationally and 5.3% among transitions originating in the target PUMA. This is despite the fact that $P(g)$ still contains 10,752 transition groups with only 294 restrictions imposed by $f(o)$ and $h(d)$. The model also fits well the origin type distribution of workers who are unemployed after the shock, particularly when distance bins are aggregated to those in earlier figures, suggesting that the counterfactual estimates of the share of employment incidence accruing to different skill/distance bin combinations are likely to be accurate. The model fit is negligibly affected by further restricting the LEHD shock sample to tract-years in which the shock is particularly likely to be driven by demand rather than supply, as judged by the existence of a single establishment in the tract that adds or lays off at least 100 workers in the chosen year.

Furthermore, the assignment model vastly outperforms a one-sided parametric conditional logit model fit to the same pre-shock CCPs $P(g|d)$. Thus, with many million observed transitions, it appears that the risk of overfitting from using a highly saturated, just identified model is far outweighed by the inability of a more parsimonious parametric model (still featuring over a hundred parameters!) to fully capture the rich multidimensional matching patterns contained in the data. The two-sided model also outperforms (though by a much smaller margin) other one-sided non-parametric forecasts that hold fixed the full set of either raw or smoothed CCPs (so $P(g)^{y,CF} = h^y(d)P^{y-1}(g|d)$), 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 employment statuses that follows substantial 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 the model estimates to simulate the welfare incidence across locations and worker skill categories of a variety of local labor demand shocks representing 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. 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 assessed among 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 or patient-doctor 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 within a very small radius of the local shock are a tiny share of the U.S. labor force competing for positions, I also find that 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 around 7.0% of the welfare gains going to existing workers in the PUMA that contains the focal 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 and spatial composition of welfare gains across workers slightly farther from the site is extremely similar across different types of demand shocks. Thus, state or national funders of local projects can afford to devolve the selection of particular local projects to fund, since such funders will internalize these 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 have jobs at risk from negative shocks, but would have been working anyway without a positive shock.

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 Monte et al. (2015), one could augment the model to allow joint residential and employment location decisions to highlight the role of commuting and house price changes in determining shock incidence, while still maintaining flexible matching among observably heterogeneous workers and positions.

References

Abowd, John M, Bryce E Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L McKinney, Marc Roemer, and Simon Woodcock, "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators," in "Producer Dynamics: New Evidence from Micro Data," University of Chicago Press, 2009, pp. 149–230.

Arcidiacono, Peter and Robert A Miller, "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity," *Econometrica*, 2011, *79* (6), 1823–1867.

Autor, David H, David Dorn, Gordon H Hanson, and Jae Song, "Trade Adjustment: Worker-level Evidence," *The Quarterly Journal of Economics*, 2014, *129* (4), 1799–1860.

Bartik, Timothy J., "Economic Development," in J. Richard Aronson and Eli Schwartz, eds., *Management Policies in Local Government Finance*, 2004.

Bartik, Timothy J and Nathan Sotherland, "Local Job Multipliers in the United States: Variation with Local Characteristics and with High-Tech Shocks," 2019.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A distributional framework for matched employer employee data," *Econometrica*, 2019, *87* (3), 699–739.

Busso, Matias, Jesse Gregory, and Patrick Kline, "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *The American Economic Review*, 2013, *103* (2), 897–947.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "The Impact of Trade on Labor Market Dynamics," Technical Report, National Bureau of Economic Research 2015.

Chen, Liang, "Compensation, Moral Hazard, and Talent Misallocation in the Market for CEOs," 2017.

Chiappori, Pierre-André and Bernard Salanié, "The Econometrics of Matching Models," *Journal of Economic Literature*, September 2016, *54* (3), 832–61.

Chiappori, Pierre-Andre, Murat Iyigun, and Yoram Weiss, "Investment in Schooling and the Marriage Market," *The American Economic Review*, 2009, *99* (5), 1689–1713.

Choo, Eugene, "Dynamic Marriage Matching: An Empirical Framework," *Econometrica*, 2015, *83* (4), 1373–1423.

_ **and Aloysius Siow**, "Who Marries Whom and Why," *Journal of Political Economy*, 2006, *114* (1), 175–201.

Decker, Colin, Elliott H Lieb, Robert J McCann, and Benjamin K Stephens, "Unique Equilibria and Substitution Effects in a Stochastic Model of the Marriage Market," *Journal of Economic Theory*, 2013, *148* (2), 778–792.

Diamond, Rebecca, "The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000," *American Economic Review*, 2016, *106* (3), 479–524.

Ekeland, Ivar, James J Heckman, and Lars Nesheim, "Identification and estimation of hedonic models," *Journal of political economy*, 2004, *112* (S1), S60–S109.

Freedman, Matthew, "Targeted Business Incentives and Local Labor Markets," *Journal of Human Resources*, 2013, *48* (2), 311–344.

Galichon, Alfred and Bernard Salanié, "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," 2015.

_ **,** _ **et al.**, "The Econometrics and Some Properties of Separable Matching Models," *American Economic Review*, 2017, *107* (5), 251–255.

Glaeser, Edward L, Joshua D Gottlieb et al., "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, 2008, *39* (1 (Spring)), 155–253.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010, *118* (3), 536–598.

Gregory, Jesse, "The Impact of Post-Katrina Rebuilding Grants on the Resettlement Choices of New Orleans Homeowners," *Work. Pap., Univ. Wisconsin–Madison*, 2013.

Heckman, James J, Rosa L Matzkin, and Lars Nesheim, "Nonparametric identification and estimation of nonadditive hedonic models," *Econometrica*, 2010, *78* (5), 1569–1591.

Hillier, Frederick S and Gerald J Lieberman, "Introduction to Operations Research. Ninth," 2010.

Hotz, V Joseph and Robert A Miller, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 1993, *60* (3), 497–529.

Hyatt, Henry R, Erika McEntarfer, Ken Ueda, and Alexandria Zhang, "Interstate Migration and Employer-to-Employer Transitions in the US: New Evidence from Administrative Records Data," *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-44*, 2016.

Kennan, John and James R Walker, "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 2011, *79* (1), 211–251.

Kline, Patrick and Enrico Moretti, "Place Based Policies with Unemployment," Technical Report, National Bureau of Economic Research 2013.

Koopmans, Tjalling C and Martin Beckmann, "Assignment Problems and the Location of Economic Activities," *Econometrica: Journal of the Econometric Society*, 1957, pp. 53–76.

LeGower, Michael and Randall Walsh, "Promise Scholarship Programs as Place-Making Policy: Evidence from School Enrollment and Housing Prices," *Journal of Urban Economics*, 2017.

Lindenlaub, Ilse, "Sorting Multidimensional Types: Theory and Application," *The Review of Economic Studies*, 2017, *84* (2), 718–789.

Manning, Alan and Barbara Petrongolo, "How Local Are Labor Markets? Evidence from a Spatial Job Search Model," *American Economic Review*, 2017.

Menzel, Konrad, "Large Matching Markets as Two-Sided Demand Systems," *Econometrica*, 2015, *83* (3), 897–941.

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg, "Commuting, Migration and Local Employment Elasticities," Technical Report, National Bureau of Economic Research 2015.

Moretti, Enrico, "Local multipliers," *The American Economic Review*, 2010, pp. 373–377.

Mourifié, Ismael and Aloysius Siow, "The Cobb Douglas Marriage Matching Function: Marriage Matching with Peer and Scale Effects," Technical Report 2017.

Neumark, David and Helen Simpson, "Place-Based Policies," Technical Report, National Bureau of Economic Research 2014.

Nimczik, Jan Sebastian, "Job Mobility Networks and Endogenous Labor Markets," 2017.

Piyapromdee, Suphanit, "The Impact of Immigration on Wages, Internal Migration and Welfare," Technical Report, Puey Ungphakorn Institute for Economic Research 2017.

Roback, Jennifer, "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 1982, *90* (6), 1257–1278.

Roth, Alvin E and Marilda Sotomayor, "Two-Sided Matching," *Handbook of Game Theory with Economic Applications*, 1992, *1*, 485–541.

Sattinger, Michael, "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature*, 1993, *31* (2), 831–880.

Schmutz, Benoît and Modibo Sidibe, "Frictional Labor Mobility," 2016.

Shapley, LS and Martin Shubik, *Game Theory in Economics: Chapter 3: the" Rules of the Games"*, Rand, 1972.

Tervio, Marko, "The Difference that CEOs Make: An Assignment Model Approach," *American Economic Review*, 2008, *98* (3), 642–68.

Weinstein, Russell, "Dynamic Responses to Labor Demand Shocks: Evidence from the Financial Industry in Delaware," *Journal of Urban Economics*, 2018.

Table 1: Specifications for Alternative Counterfactual Labor Demand Shocks

Spec. No.	Number of Jobs (or % of Tract's Jobs)	Firm Avg. Earn. Quartile	Firm Size Quartile	Industry Supersector	Shock Type
1	500	2	1	Information	Stimulus
2	500	2	4	Information	Stimulus
3	500	4	1	Information	Stimulus
4	500	4	4	Information	Stimulus
5	500	2	1	Manufacturing	Stimulus
6	500	2	4	Manufacturing	Stimulus
7	500	4	1	Manufacturing	Stimulus
8	500	4	4	Manufacturing	Stimulus
9	500	2	1	Trade/Trans./Utilities	Stimulus
10	500	2	4	Trade/Trans./Utilities	Stimulus
11	500	4	1	Trade/Trans./Utilities	Stimulus
12	500	4	4	Trade/Trans./Utilities	Stimulus
13	500	2	1	Other Services	Stimulus
14	500	2	4	Other Services	Stimulus
15	500	4	1	Other Services	Stimulus
16	500	4	4	Other Services	Stimulus
17	500	2	1	Education & Health	Stimulus
18	500	2	4	Education & Health	Stimulus
19	500	4	1	Education & Health	Stimulus
20	500	4	4	Education & Health	Stimulus
21	500	2	1	Leisure & Hospitality	Stimulus
22	500	2	4	Leisure & Hospitality	Stimulus
23	500	4	1	Leisure & Hospitality	Stimulus
24	500	4	4	Leisure & Hospitality	Stimulus
25	500	2	1	Government	Stimulus
26	500	2	4	Government	Stimulus
27	500	4	1	Government	Stimulus
28	500	4	4	Government	Stimulus
29	500	2	1	Construction	Stimulus
30	500	2	4	Construction	Stimulus
31	500	4	1	Construction	Stimulus
32	500	4	4	Construction	Stimulus
33	500	4	1	Information	Restr. Stim.
34	500	2	4	Manufacturing	Restr. Stim.
35	500	2	1	Trade/Trans./Utilities	Restr. Stim.
36	25%	All	All	All	Disaster
37	50%	All	All	All	Disaster
38	100%	All	All	All	Disaster

Table 2: Summary Statistics Describing Heterogeneity in the Spacial Scope of Labor Markets by Origin Worker Characteristics and Destination Establishment Characteristics

Panel A: By Origin Worker Unemployment or Earnings Category

		Share of All Transitions				Share of Transitions to New Jobs		
Worker Subpop.	# of Obs.	Unemp. to Unemp.	Emp. to Unemp.	Stay at Same Job	Move to New Job	Same PUMA	New PUMA, Same State	New State
All Workers	23485000	0.042	0.043	0.703	0.213	0.320	0.600	0.080
Young (<25) Unemp.	1018000	0.154	0.000	0.000	0.846	0.267	0.636	0.096
Older ($\geq$ 25) Unemp.	1747000	0.469	0.000	0.000	0.531	0.287	0.600	0.113
1st Earn. Quart.	4947000	0.000	0.097	0.667	0.236	0.343	0.604	0.053
2nd Earn. Quart.	5028000	0.000	0.052	0.784	0.164	0.355	0.581	0.064
3rd Earn. Quart.	5121000	0.000	0.030	0.849	0.121	0.350	0.578	0.073
4th Earn. Quart.	5647000	0.000	0.020	0.873	0.107	0.331	0.583	0.086

Panel B: By Destination Establishment Pay Quartile and Size Quartile

		Share of All Transitions				Share of Transitions to New Jobs		
Estab. Subpop.	# of Obs.	Unemp. to Unemp.	Emp. to Unemp.	Stay at Same Job	Move to New Job	Same PUMA	New PUMA, Same State	New State
1st Q. Avg. Earn.	5932000	0.000	0.000	0.668	0.332	0.323	0.603	0.075
2nd Q. Avg. Earn.	5362000	0.000	0.000	0.783	0.217	0.334	0.590	0.077
3rd Q. Avg. Earn.	4845000	0.000	0.000	0.815	0.185	0.331	0.591	0.078
4th Q. Avg. Earn.	5382000	0.000	0.000	0.819	0.181	0.292	0.612	0.096
1st Q. Size	2374000	0.000	0.000	0.682	0.318	0.284	0.446	0.271
2nd Q. Size	2063000	0.000	0.000	0.722	0.278	0.344	0.592	0.065
3rd Q. Size	2091000	0.000	0.000	0.727	0.273	0.326	0.625	0.049
4th Q. Size	15003000	0.000	0.000	0.793	0.207	0.324	0.634	0.043

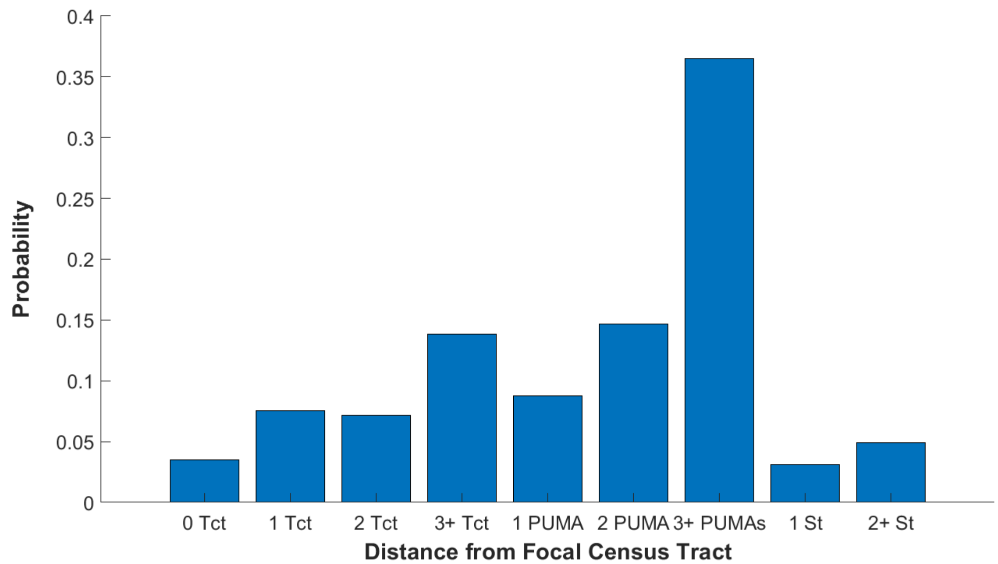
Panel C: By Destination Establishment Industry

		Share of All Transitions				Share of Transitions to New Jobs		
Estab. Industry	# of Obs.	Unemp. to Unemp.	Emp. to Unemp.	Stay at Same Job	Move to New Job	Same PUMA	New PUMA, Same State	New State
Nat. Resources	369900	0.000	0.000	0.692	0.308	0.496	0.406	0.097
Construction	974700	0.000	0.000	0.741	0.259	0.319	0.586	0.095
Manufacturing	1929200	0.000	0.000	0.840	0.160	0.398	0.530	0.072
Whole/Retail/Trans.	4382000	0.000	0.000	0.764	0.236	0.280	0.635	0.085
Information	539100	0.000	0.000	0.799	0.201	0.318	0.599	0.082
Financial Activities	1287800	0.000	0.000	0.769	0.231	0.285	0.648	0.067
Prof. Bus. Services	3020000	0.000	0.000	0.692	0.308	0.275	0.642	0.083
Ed. Health	4859000	0.000	0.000	0.815	0.185	0.363	0.566	0.071
Leis. & Hosp.	2269000	0.000	0.000	0.670	0.330	0.328	0.583	0.088
Oth. Serv.	841300	0.000	0.000	0.744	0.256	0.330	0.604	0.066
Government	1059900	0.000	0.000	0.910	0.090	0.393	0.550	0.057

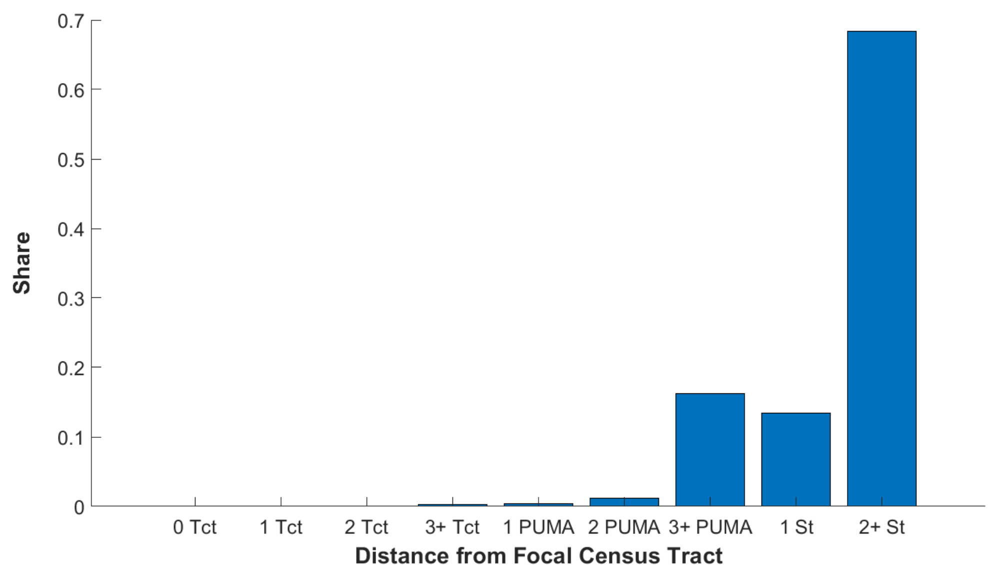
Notes: "1st Q. Avg. Earn": 1st quartile of the establishment-level worker avg. annual earnings distribution. "1st Q. Size": 1st quartile of the establishment-level employment distribution.

Figure 1: Key Distributions

(a) Empirical Distribution of 2010-2011 Job Transitions



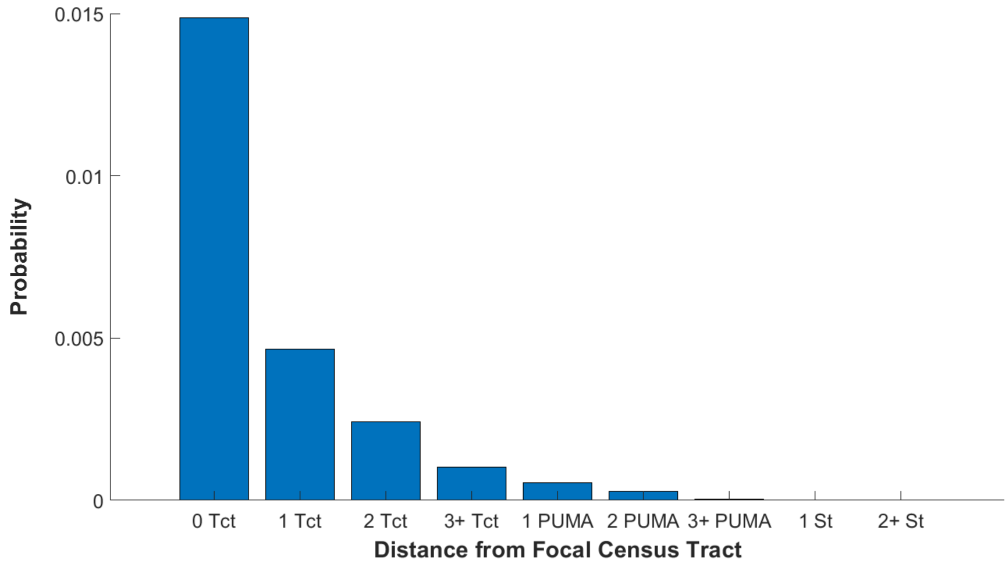
(b) Distribution of the Distance Between Workers' Origin Position and the Census Tract Targeted by the Simulated Stimulus Package: Average across All Simulated Stimuli



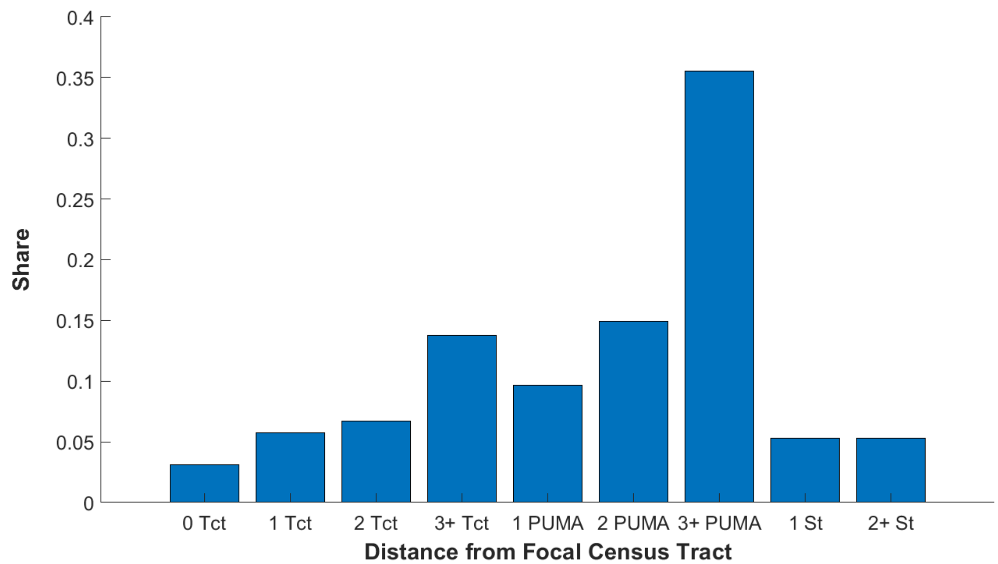
Notes: The bar heights in Figure 1a capture the shares of all worker transitions between dominant positions in 2010 and 2011 in which the geographic distance between these positions' establishments fell into the distance bins indicated by the bar labels. The bar heights in Figure 1b capture the shares of all workers for whom the geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the labeled distance bins (computed separately for each target tract, then averaged across all 500 target tracts). "0/1/2/3+ Tct" indicates that the two establishments (or, for figure 1b, the establishment and the targeted tract) 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: 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



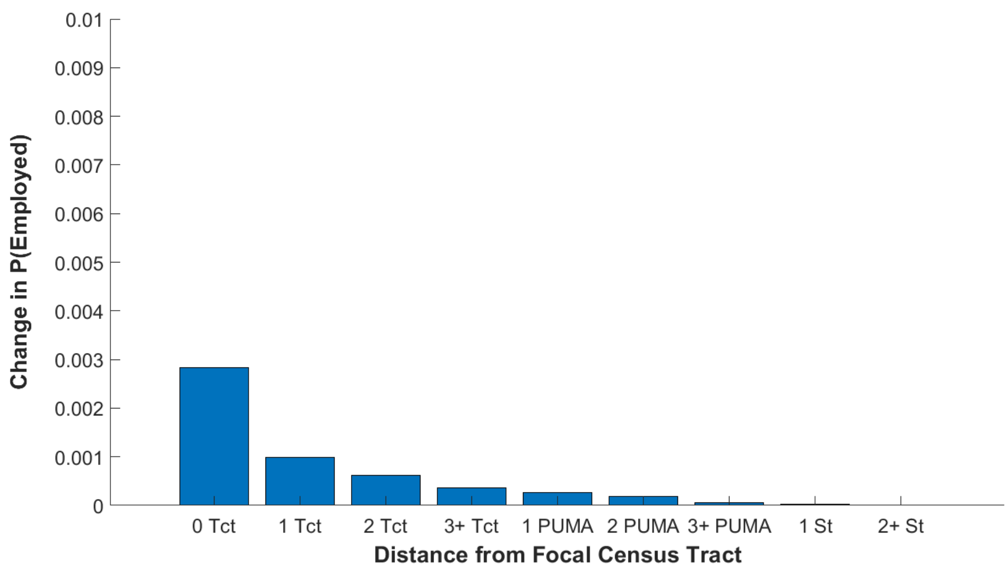
(b) Share of All Stimulus Jobs



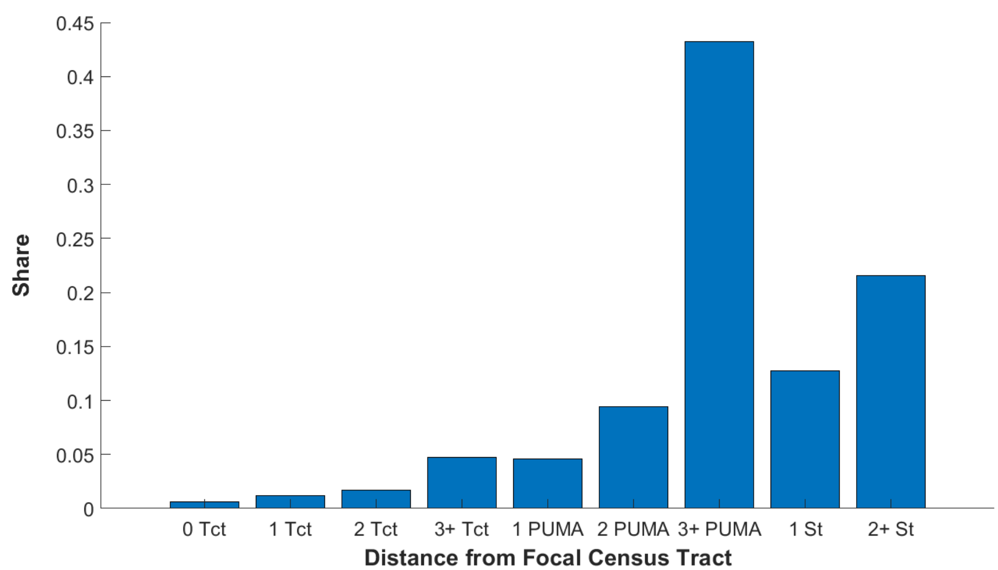
Notes: The bar heights in Figure 2a capture the average probability of obtaining a stimulus job among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels. These probabilities average across different initial earnings categories and across stimulus packages featuring different firm compositions. Figure 2b 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 3: 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)



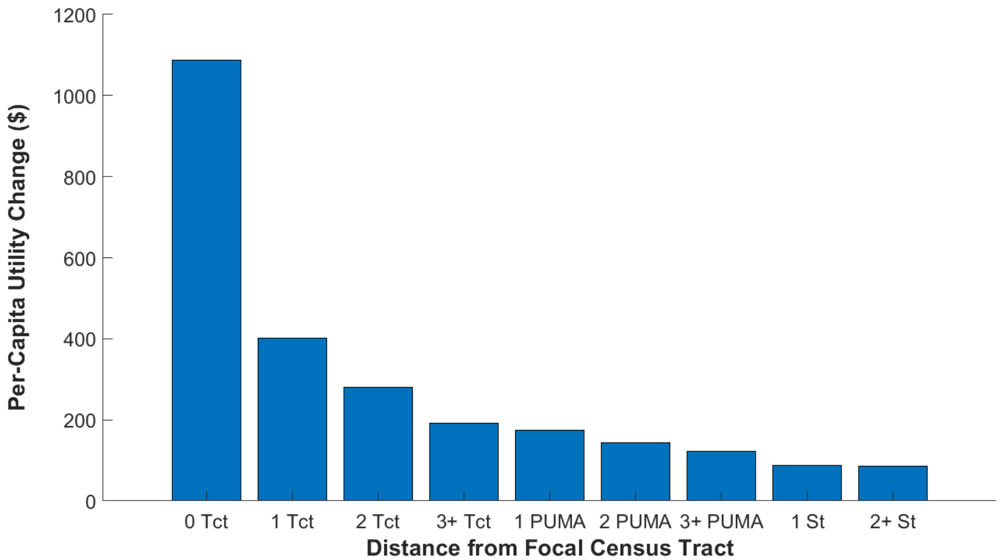
(b) Share of All Additional Employment



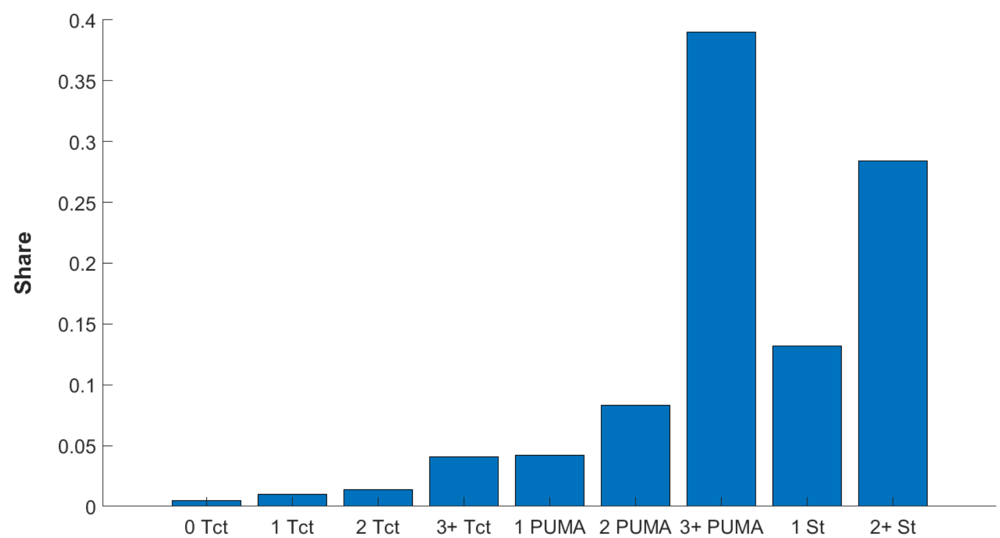
Notes: The bar heights in Figure 3a capture the average change in employment probability among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels. These probabilities average across different initial earnings categories and across stimulus packages featuring different firm compositions. Figure 3b 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 4: Expected Welfare Changes and Share of Total Welfare Gains by Distance From Focal Tract: Average across All Simulated Stimuli

(a) Expected Welfare Changes ($)



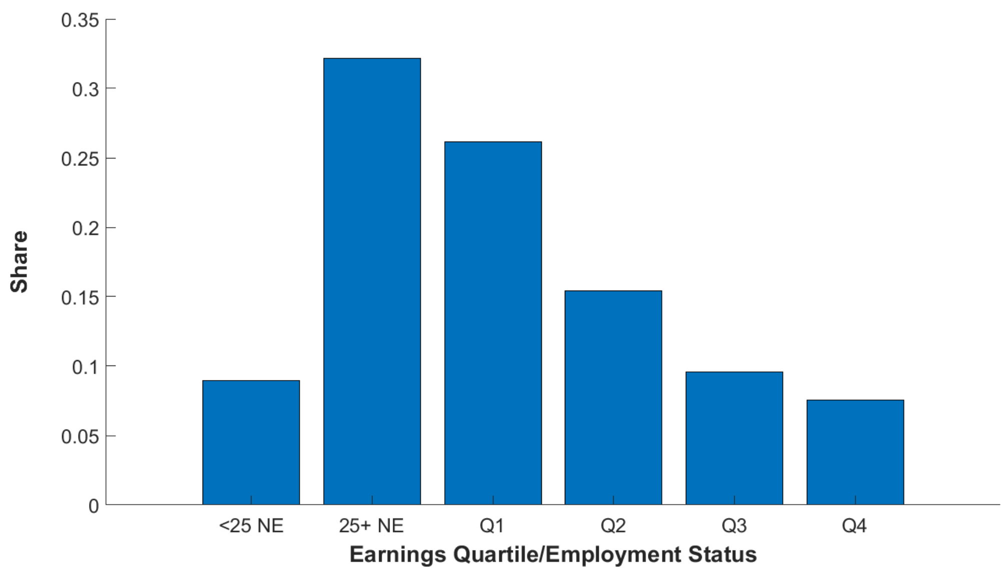
(b) Share of Total Welfare Gains



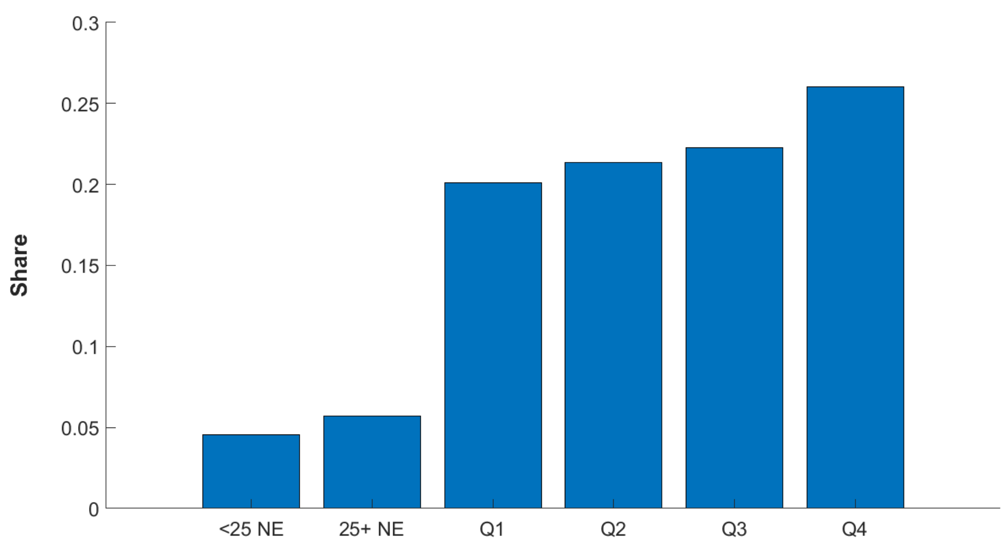
Notes: The bar heights in Figure 4a capture the average welfare gain (scaled in $ of annual earnings) among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins indicated by the bar labels. Averages are taken across different initial earnings categories and across stimulus packages featuring different firm compositions. Figure 4b 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 5: 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



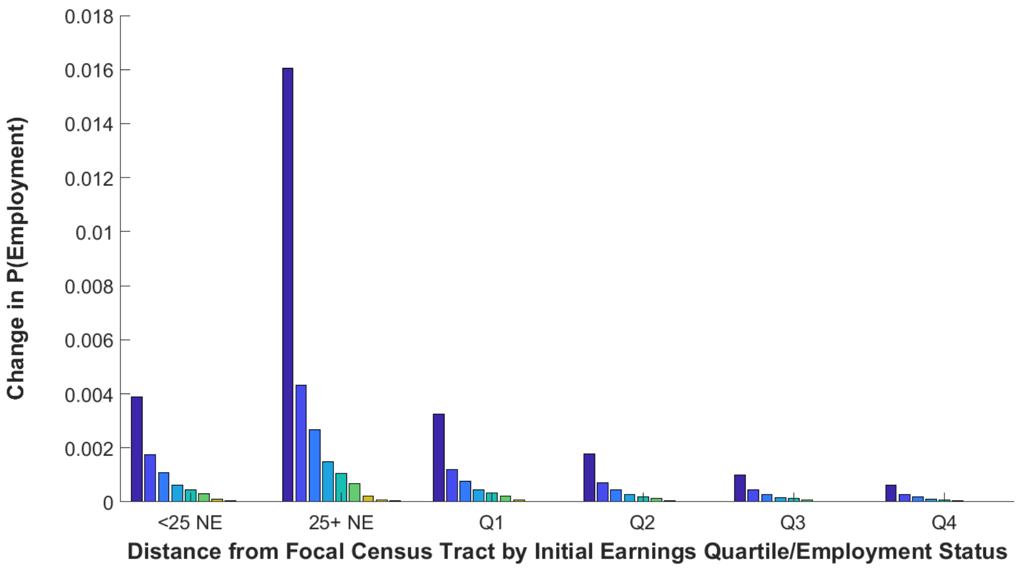
(b) Share of Total Utility Gains



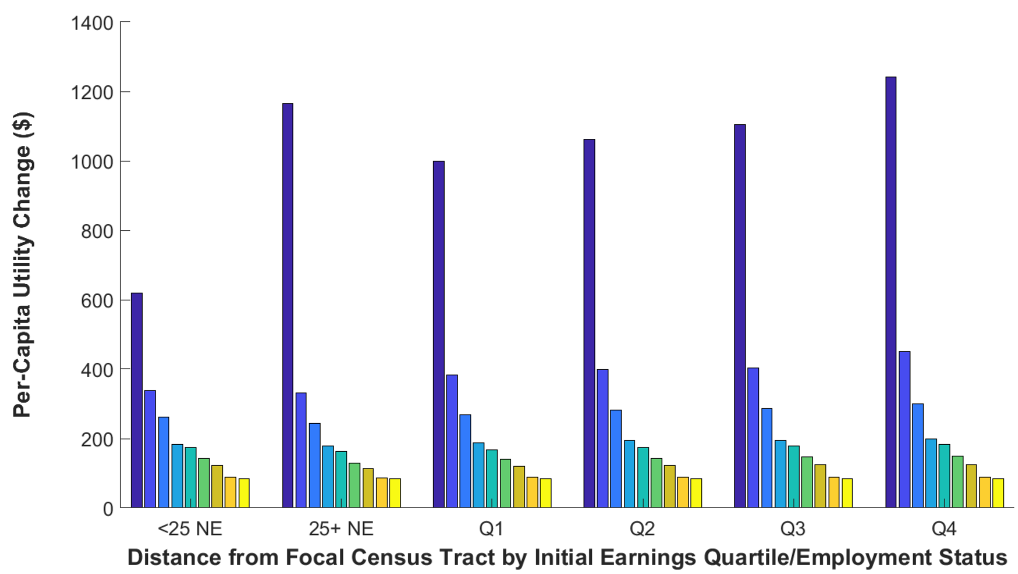
Notes: The bar heights in Figure 5a and 5b capture the average share of all employment gains and of all welfare gains, respectively, among workers whose origin employment status fell into the employment status/earnings quartiles indicated by the bar labels. Averages are taken across different bins capturing 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 6: 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)



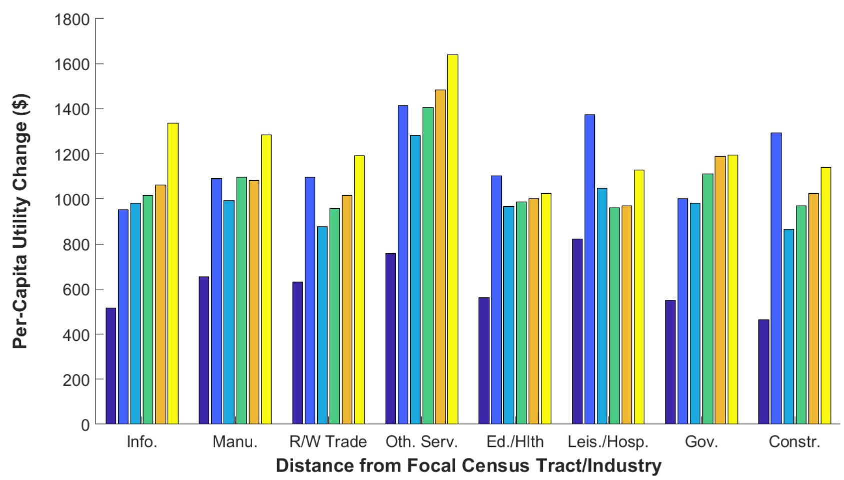
(b) Change in Expected Utility



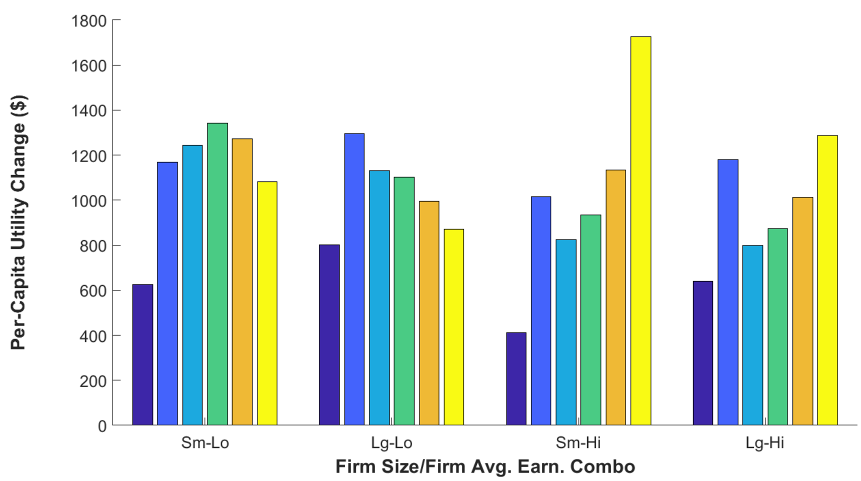
Notes: The bar heights within a particular group in Figures 6a and 5b capture the average change in employment probability and share of all additional employment, respectively, among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins defined in Figure 2. 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 7: 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

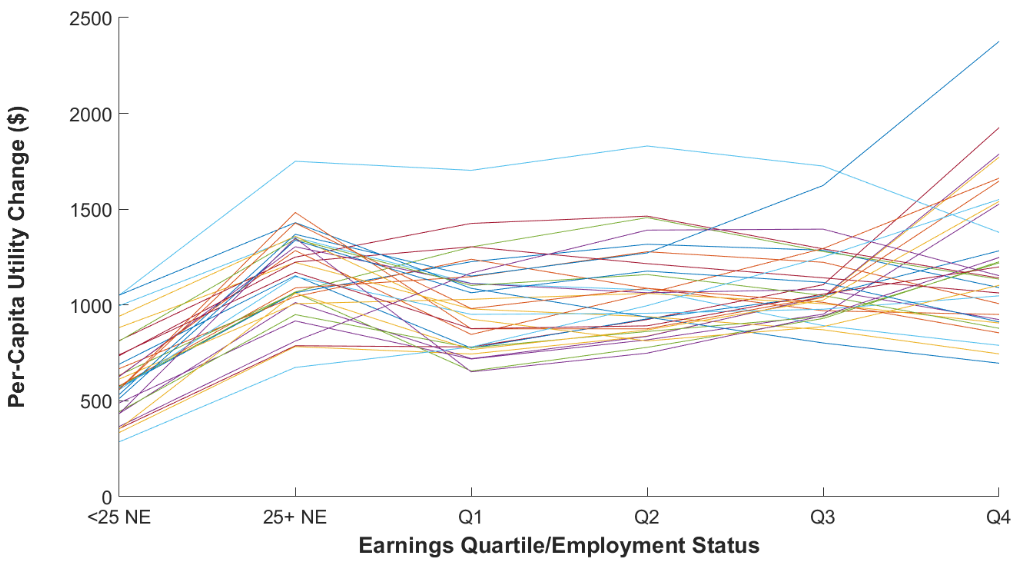


(b) Expected Welfare Gain by Firm Size Quartile/Firm Avg. Pay Quartile Combination



Notes: The bar heights within a particular group in Figures 7a and 7b 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 5. Each group of bars displays average outcomes among simulated stimulus packages featuring positions within the particular industry supersector (in Figure 7a) or particular firm size/firm avg. earnings quartile combination (in Figure 7b) 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 7b) or different firm size/firm pay compositions (in Figure 7a), 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. "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.

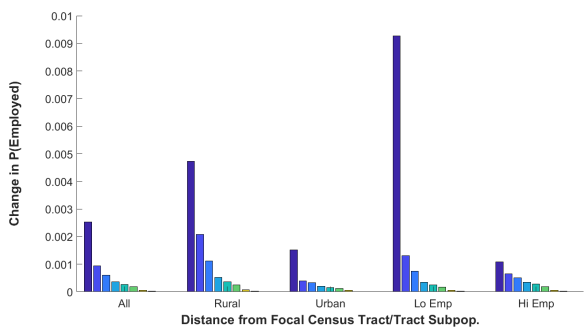
Figure 8: Expected Utility Changes Among Workers Originally Working in the Targeted Tract by Initial Earnings/Employment Status: All Stimulus Packages



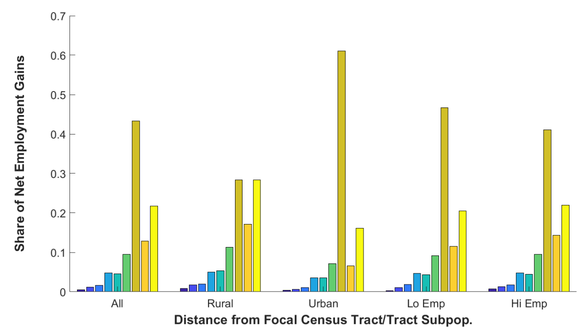
Notes: Each line traces the expected welfare gain generated by a stimulus package featuring 500 positions among firms with a particular supersector/firm size quartile/firm pay quartile combination across alternative origin nonemployment or earnings quartile categories. 32 different lines corresponding 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 9: Heterogeneity in Several Incidence Measures by Distance from Focal Tract Across Focal Tracts of Varying Population and Employment Size

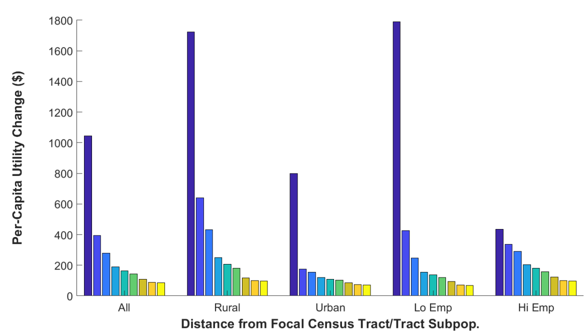
(a) Change in P(Employed)



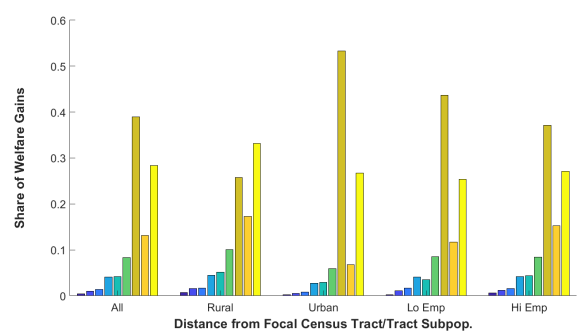
(b) Share of Employment Gains



(c) Avg. Welfare Gain ($)



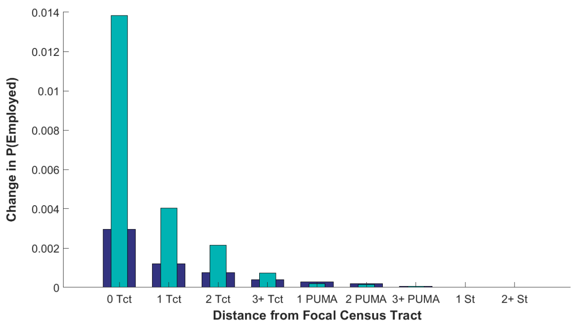
(d) Share of Welfare Gains



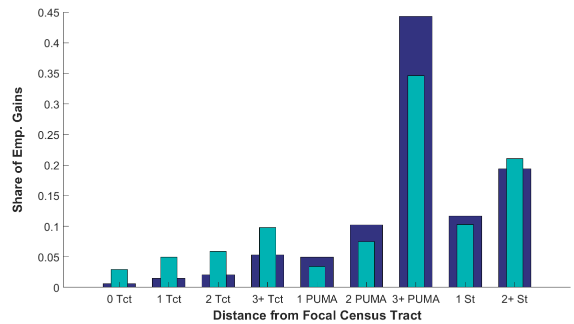
Notes: The bar heights within a particular group in Figures 9a-9d capture the average measure of stimulus incidence associated with the chosen figure from a 500 person stimulus package among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance bins defined in Figure 2. 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 10: 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)

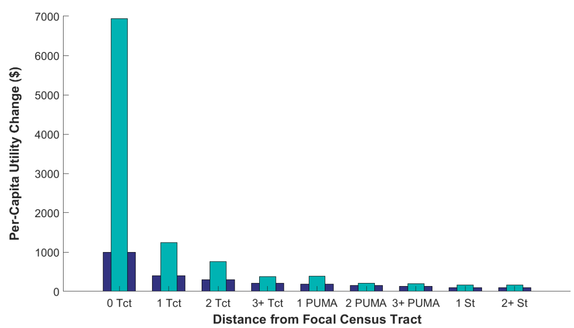
(a) Change in P(Employed)



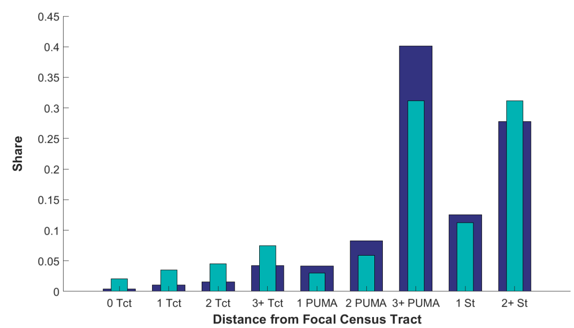
(b) Share of Employment Gains



(c) Avg. Welfare Gain ($)



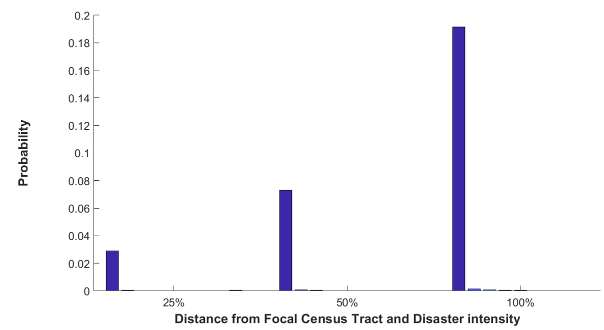
(d) Share of Welfare Gains



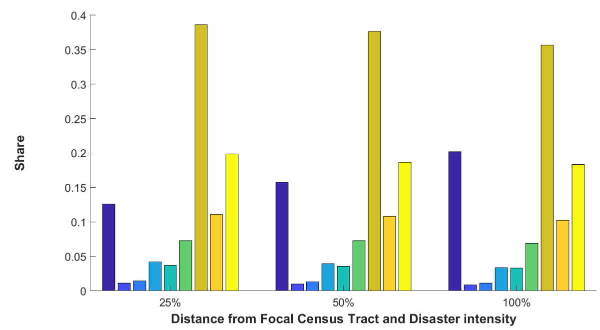
Notes: The bar heights capture the average measure of stimulus incidence associated with the chosen figure from a 500 person stimulus package among workers whose geographic distance between their origin establishments and the census tract receiving the simulated stimulus package fell into the distance 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 11: Employment and Welfare Incidence from a Natural Disaster by Distance From Focal Tract and Severity of the Disaster (25%/50%/100% Jobs Lost)

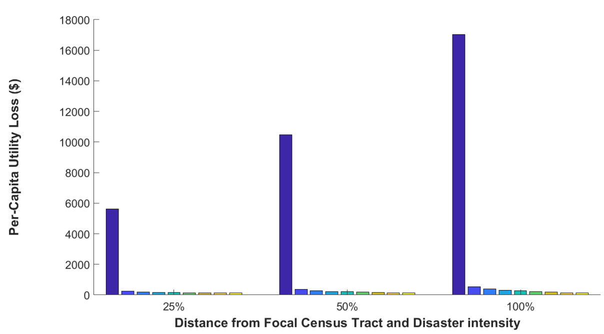
(a) Change in P(Unemployed)



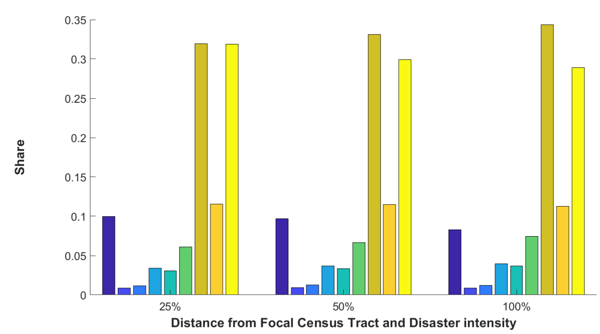
(b) Share of Employment Losses



(c) Expected Welfare Loss



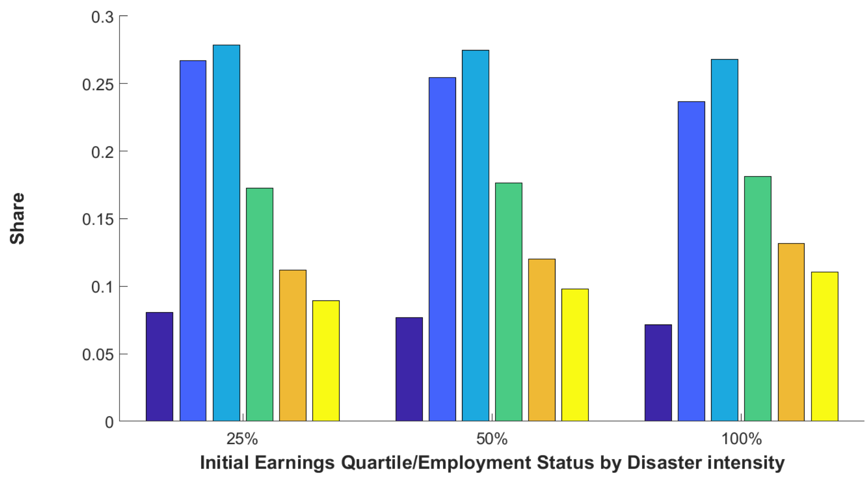
(d) Share of Total Welfare Losses



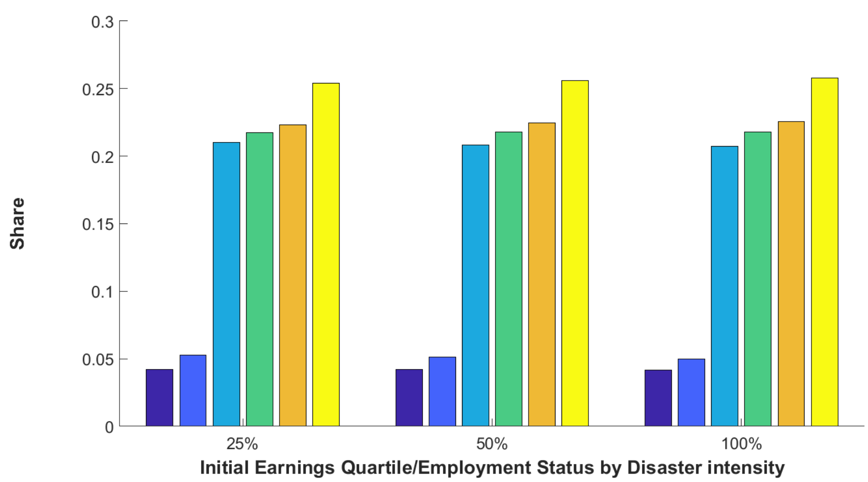
Notes: The bar heights within a particular group in Figures 11a-11d capture the average value of the incidence measure associated with the figure from a set of simulated natural disasters among workers whose geographic distance between their origin establishments and the census tract experiencing the disaster fell into the distance bins defined in Figure 2. 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 12: 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)

(a) Share of Additional Nonemployment



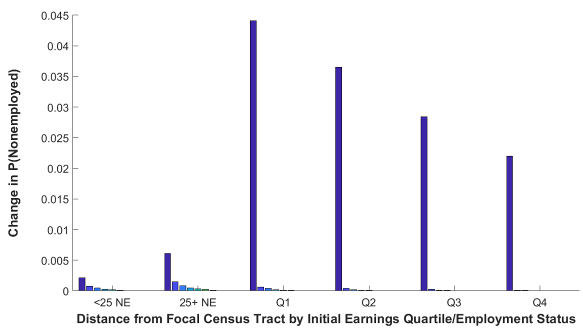
(b) Share of Total Welfare Losses



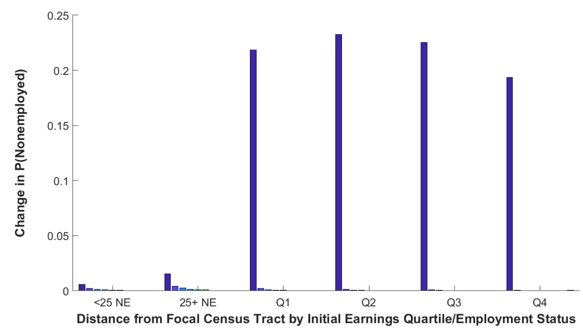
Notes: The bar heights within a particular group in Figures 12a and 12b capture the average share of additional nonemployment and welfare losses, respectively, from a set of simulated natural disasters among workers whose origin employment status/earnings quartile fell into the bins defined in Figure 5. 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 13: 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

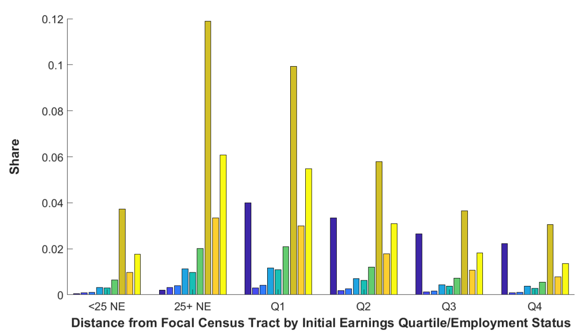
(a) Change in P(Unemployed), 25% of Jobs Lost



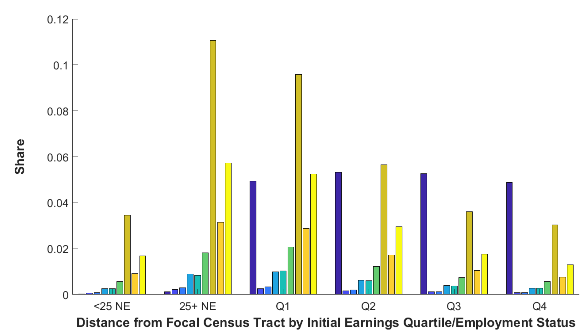
(b) Change in P(Unemployed), 100% of Jobs Lost



(c) Share of Total Employment Losses, 25% of Jobs Lost



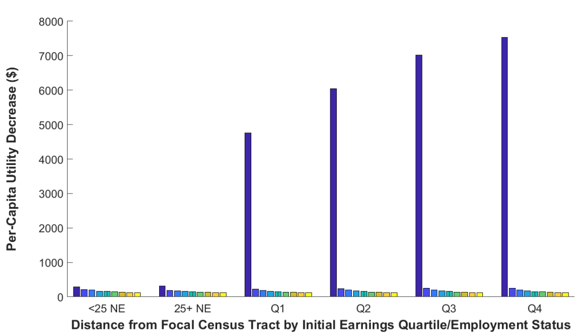
(d) Share of Total Employment Losses, 100% of Jobs Lost



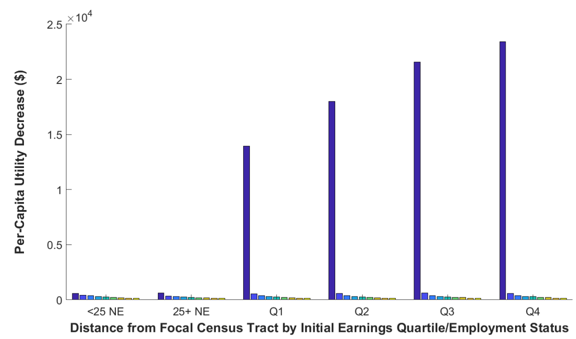
Notes: The bar heights within a particular group in Figures 13a-13d capture either the average change in unemployment probability or the share of all employment losses (depending on the figure) from a set of simulated natural disasters among workers whose geographic distance between their origin establishments and the census tract experiencing the disaster fell into the distance bins defined in Figure 2. In Figures 13a and 13c, 25% of jobs in the targeted census tract are eliminated, while in Figures 13b and 13d 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 14: 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

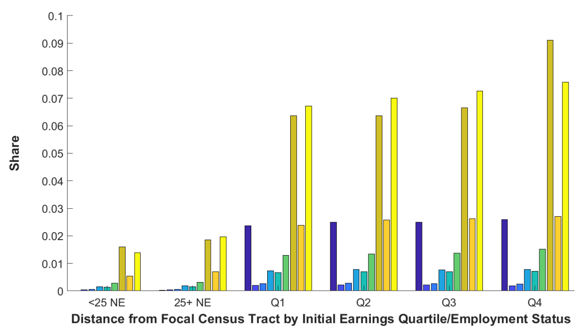
(a) Expected Welfare Loss, 25% of Jobs Lost



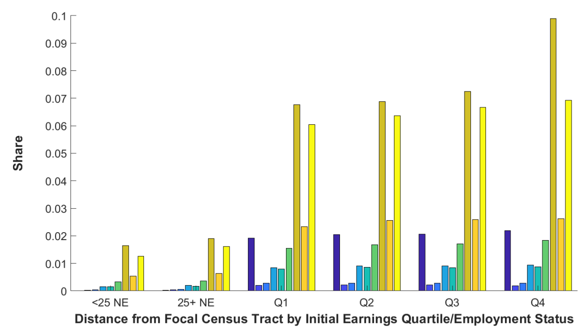
(b) Expected Welfare Loss, 100% of Jobs Lost



(c) Share of Total Welfare Losses, 25% of Jobs Lost



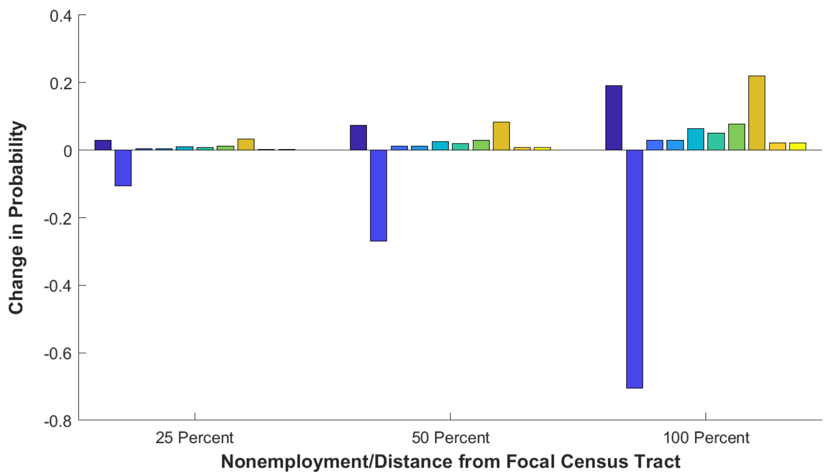
(d) Share of Total Welfare Losses, 100% of Jobs Lost



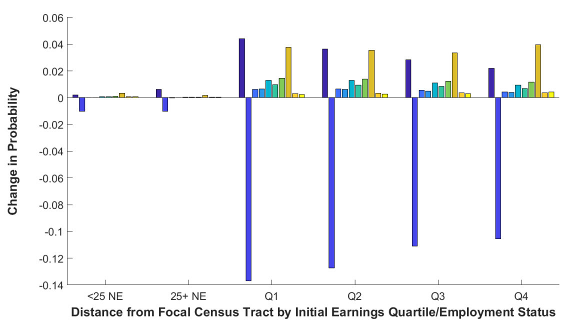
Notes: The bar heights within a particular group in Figures 13a-13d capture either the average welfare loss (scaled in $ of annual earnings) or the share of all welfare losses (depending on the figure) from a set of simulated natural disasters among workers whose geographic distance between their origin establishments and the census tract experiencing the disaster fell into the distance bins defined in Figure 2. In Figures 14a and 14c, 25% of jobs in the targeted census tract are eliminated, while in Figures 14b and 14d 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 15: 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)

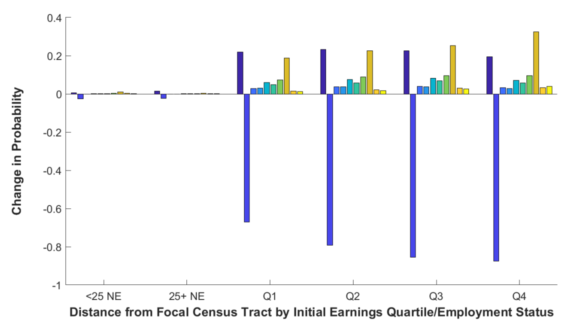
(a) 25/50/100% of Jobs Lost, at Different Distances from the Focal Tract (Averaging Across Initial Earnings/Employment Statuses)



(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 15a-15c 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 2 (or become/remain unemployed, the leftmost bar in each group). In Figure 15a, 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 15b and 15c, 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.

Online Appendix

A1 Proof of Proposition A1

Proposition A1:

Let $|o|$ 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 $o(g)$). In addition, let $f(o)$ denote the share of all workers assigned to origin type o, so that $|o| = f(o)I$. Further, define C_o as the mean value of $e^{-\frac{r_i}{\sigma}}$ for a given origin group o. Define $S_{g|o,k}$ as 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 define $\overline{S}_{g|o,d}$ to be the mean of $S_{g|o,k}$ among all k assigned to destination type d. Suppose the following assumptions hold:

$$\text{Assumption 1: } \frac{1}{|g_k|}\sum_{i:g(i,k)=g} e^{-\frac{r_i}{\sigma}} \approx \frac{1}{|o|}\sum_{i:o(i)=o(g)} e^{-\frac{r_i}{\sigma}} = C_{o(g)} \;\forall (g,k) \tag{14}$$

$$\text{Assumption 2: } S_{g|o,k} \approx \overline{S}_{g|o,d} \;\forall\, k,\; \forall\, g \tag{15}$$

Then the equilibrium aggregate group-level choice probabilities can be written as follows:

$$P(g|d) = \frac{e^{\frac{\theta_g}{\sigma}}\overline{S}_{g|o,d} f(o) C_o}{\sum_{o'\in\mathcal{O}}\sum_{g'\in(o,d)} e^{\frac{\theta_{g'}}{\sigma}}\overline{S}_{g'|o',d} f(o') C_{o'}} \tag{16}$$

Proof: First, note that the law of total probability implies:

$$\begin{aligned} P(g|d) &= \sum_{k\in d} P(g|d,k)P(k|d) = \frac{1}{|d|}\sum_{k\in d} P(g|k) = \frac{1}{|d|}\sum_{k\in d}\sum_{i:g(i,k)=g} P(i|k) \\ &= \frac{1}{|d|}\sum_{k\in d}\sum_{i:g(i,k)=g} \frac{e^{\frac{\theta_g - r_i}{\sigma}}}{\sum_{i'\in\mathcal{I}} e^{\frac{\theta_{g'} - r_{i'}}{\sigma}}} = \frac{1}{|d|}\sum_{k\in d} \frac{(e^{\frac{\theta_g}{\sigma}})(\sum_{i:g(i,k)=g} e^{\frac{-r_i}{\sigma}})}{\sum_{i'\in\mathcal{I}} e^{\frac{\theta_{g'} - r_{i'}}{\sigma}}}, \end{aligned} \tag{17}$$

where $|d|$ captures the number of positions k assigned to destination type d.

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.[50] In other words, the payoffs that workers in the

[50] 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.

same skill class require in equilibrium will not differ systematically across establishments within a small local area. This becomes a better approximation as more characteristics (such as occupation or education) are used to define an origin type $o(i)$.

Assumption 2 imposes that the share of potential stayers vs. new hires among workers from each origin type o is common across establishments within destination type d. In the chosen context, this means that establishments in the same geographic area, industry supersector, and establishment size and average pay categories have roughly the same number and skill composition of employees.

These assumptions are necessary because the aggregate mean of a non-linear function of a random variable (in this case $e^{\frac{-r_i}{\sigma}}$) 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).

Note first that Assumption 2 implies that $|g_k| \equiv S_{g|o,k} f(o(g))I \approx \overline{S}_{g|o,d} f(o(g))I$. Thus, Assumptions 1 and 2 together imply:

$$\sum_{i:g(i,k)=g} e^{-\frac{r_i}{\sigma}} \approx \overline{S}_{g|o(g),d(g)} f(o(g))(I)C_{o(g)}. \tag{18}$$

Applying this result to the last expression in (17), one obtains:

$$\begin{aligned}
P(g|d) &= \sum_{k\in d}(\frac{1}{|d|})\frac{e^{\frac{\theta_g}{\sigma}}\sum_{i:g(i,k)=g} e^{-\frac{r_i}{\sigma}}}{\sum_{i'\in\mathcal{I}} e^{\frac{\theta_{g'}-r_{i'}}{\sigma}}} = \sum_{k\in d}(\frac{1}{|d|})\frac{e^{\frac{\theta_g}{\sigma}}\sum_{i:g(i,k)=g} e^{-\frac{r_i}{\sigma}}}{\sum_{o'\in\mathcal{O}}\sum_{g'\in(o,d)}\sum_{i':g(i',k)=g'} e^{\frac{\theta_{g'}-r_{i'}}{\sigma}}} \\
&= \sum_{k\in d}(\frac{1}{|d|})\frac{e^{\frac{\theta_g}{\sigma}}\overline{S}_{g|o,d} f(o)(I)C_o}{\sum_{o'\in\mathcal{O}}\sum_{g'\in(o,d)} e^{\frac{\theta_{g'}}{\sigma}}\overline{S}_{g'|o',d} f(o')(I)C_{o'}} \\
&= \frac{e^{\frac{\theta_g}{\sigma}}\overline{S}_{g|o,d} f(o)(I)C_o}{\sum_{o'\in\mathcal{O}}\sum_{g'\in(o,d)} e^{\frac{\theta_{g'}}{\sigma}}\overline{S}_{g'|o',d} f(o')(I)C_{o'}}\sum_{k\in d}(\frac{1}{|d|}) = \frac{e^{\frac{\theta_g}{\sigma}}\overline{S}_{g|o,d} f(o)C_o}{\sum_{o'\in\mathcal{O}}\sum_{g'\in(o,d)} e^{\frac{\theta_{g'}}{\sigma}}\overline{S}_{g'|o',d} f(o')C_{o'}}
\end{aligned} \tag{19}$$

This concludes the proof.

A2 Proof of Proposition 1

Proposition 1:

Define the set $\Theta^{D-in-D} \equiv \{\frac{(\theta_g-\theta_{g'})-(\theta_{g''}-\theta_{g'''})}{\sigma}\,\forall\,(g,g',g'',g''') : o(g)=o(g''), o(g')=o(g'''), d(g)=d(g'), d(g'')=d(g''')\}$. *Given knowledge of* Θ^{D-in-D}, *a set* $\tilde{\Theta}=\{\tilde{\theta}_g\,\forall\,g\in\mathcal{G}\}$ *can be constructed such that the unique group level assignment* $P^{CF}(g)$ *that satisfies the market-clearing conditions*

(13) using $\theta_g^{CF} = \tilde{\theta}_g \; \forall \; g$ and arbitrary marginal PMFs for origin and destination types $f^{CF}(*)$ and $g^{CF}(*)$ will also satisfy the corresponding market-clearing conditions using $\theta_g^{CF} = \theta_g \; \forall \; g \in \mathcal{G}$ and the same PMFs $f^{CF}(*)$ and $g^{CF}(*)$. Furthermore, denote by $\tilde{\mathbf{C}}^{\mathbf{CF}} \equiv \{\tilde{C}_1^{CF}, \ldots, \tilde{C}_O^{CF}\}$ and $\mathbf{C}^{\mathbf{CF}} \equiv \{C_1^{CF}, \ldots, C_O^{CF}\}$ the utility vectors that clear the market using $\theta_g^{CF} = \tilde{\theta}_g$ and using $\theta_g^{CF} = \theta_g$, respectively. Then $\tilde{\mathbf{C}}^{\mathbf{CF}}$ will satisfy $\tilde{C}_o^{CF} = C_o^{CF} e^{\frac{-\Delta_o}{\sigma}} \; \forall \; o \in \mathcal{O}$ for some set of origin type-specific constants $\{\Delta_o : o \in [1, O]\}$ that is invariant to the choices of $f^{CF}(*)$ and $g^{CF}(*)$.

Proof: I prove Proposition 1 by construction.

Let $z(i,k) = 1(m(j(i)) = m(k))$ represent an indicator that takes on the value of 1 if the firms associated with positions $j(i)$ 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 $z(i,k)$. 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 $\Theta^{D-in-D} = \{\frac{(\theta_g - \theta_{g'}) - (\theta_{g''} - \theta_{g'''})}{\sigma} \; \forall \; (g, g', g'', g''')\}$ 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 set of alternative group-level joint surplus values $\tilde{\Theta} = \{\tilde{\theta}_g\}$ as follows:

$$\tilde{\theta}_{g'} = 0 \; \forall \; g' : (o(g') = 1 \text{ and/or } d(g') = 1) \text{ and } z(g') = 0 \tag{20}$$

$$\tilde{\theta}_{g'} = \frac{(\theta_{g'} - \theta_{g(1,d(g'),0)}) - (\theta_{g(o(g'),1,0)} - \theta_{g(1,1,0)})}{\sigma} \; \forall \; g' : (d(g') \neq \text{ and } o(g') \neq 1) \text{ and/or } z(g') \neq 0 \tag{21}$$

Under the definitions in (20) and (21), we have:

$$\frac{(\tilde{\theta}_g - \tilde{\theta}_{g'}) - (\tilde{\theta}_{g''} - \tilde{\theta}_{g'''})}{\sigma} = \frac{(\theta_g - \theta_{g'}) - (\theta_{g''} - \theta_{g'''})}{\sigma}$$
$$\forall \; (g, g', g'', g''') : o(g) = o(g''), o(g') = o(g'''), d(g) = d(g'), d(g'') = d(g''') \tag{22}$$

Thus, the appropriate difference-in-differences using elements of $\tilde{\Theta}$ match their analogues among the true surpluses in Θ^{D-in-D}, so that all the information about Θ in the identified set Θ^{D-in-D} is retained. And unlike the true set Θ, the construction of $\tilde{\Theta}$ only requires knowledge of Θ^{D-in-D}.

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

$$\tilde{\theta}_g = \theta_g + \Delta^1_{o(g)} + \Delta^2_{d(g)} \; \forall \; g \in \mathcal{G}, \text{ where} \tag{23}$$

$$\Delta^1_{o(g)} = \theta_{g(o(g),1,0)} - \theta_{g(1,1,0)} \quad \text{and} \quad \Delta^2_{d(g)} = \theta_{g(1,d(g),0)} \tag{24}$$

where $\mathcal{G}$ is the set of all possible transition groups. In other words, each alternative surplus $\tilde{\theta}_g$ equals the true surplus θ_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 $\tilde{P}^{CF}(*) \equiv P^{CF}(*|\tilde{\Theta}, \tilde{C}^{CF}_2, \ldots, \tilde{C}^{CF}_O))$ represent the assignment that results from combining arbitrary marginals $f^{CF}(o)$ and $h^{CF}(d)$ with $\tilde{\Theta}$. $\tilde{\mathbf{C}}^{\mathbf{CF}} = [1, \tilde{C}^{CF}_2 \ldots \tilde{C}^{CF}_O]$ 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 (10) along with the elements of $\tilde{\Theta}$, f^{CF} and $\overline{S}^{CF}_{g'|o(g'),d}$:

$$\sum_{d \in \mathcal{D}} h^{CF}(d) (\sum_{g:o(g)=2} P^{CF}(g|d, \tilde{\Theta}, \tilde{\mathbf{C}}^{\mathbf{CF}})) = f^{CF}(2)$$
$$\vdots$$
$$\sum_{d \in \mathcal{D}} h^{CF}(d) (\sum_{g:o(g)=O} P^{CF}(g|d, \tilde{\Theta}, \tilde{\mathbf{C}}^{\mathbf{CF}})) = f^{CF}(O) \tag{25}$$

I wish to show that $\tilde{P}^{CF}(*) \equiv P^{CF}(*|\tilde{\Theta}, \tilde{\mathbf{C}}^{\mathbf{CF}})$ will be identical to the alternative unique counterfactual equilibrium assignment $P^{CF}(*|\Theta, \mathbf{C}^{\mathbf{CF}})$ that combines the same arbitrary marginal distributions $f^{CF}(o)$ and $h^{CF}(d)$ with the set Θ instead of $\tilde{\Theta}$. Here, $\mathbf{C}^{\mathbf{CF}} = [1, C^{CF}_2 \ldots C^{CF}_O]$ denotes a vector of o-type-specific mean exponentiated utility values that clears the market by satisfying the following alternative excess demand equations:[51]

$$\sum_{d \in \mathcal{D}} h^{CF}(d) (\sum_{g:o(g)=2} P^{CF}(g|d, \Theta, \mathbf{C}^{\mathbf{CF}})) = f^{CF}(2)$$
$$\vdots$$
$$\sum_{d \in \mathcal{D}} h^{CF}(d) (\sum_{g:o(g)=O} P^{CF}(g|d, \Theta, \mathbf{C}^{\mathbf{CF}})) = f^{CF}(O) \tag{26}$$

Since all other terms are shared by the systems (25) and (26), it suffices to show that $P^{CF}(g|d, \tilde{\Theta}, \tilde{\mathbf{C}}^{\mathbf{CF}}) = P^{CF}(g|d, \Theta, \mathbf{C}^{\mathbf{CF}}))\ \forall\ g \in \mathcal{G}$ for some vector $\mathbf{C}^{\mathbf{CF}}$. Consider the following vector $\mathbf{C}^{\mathbf{CF}}$:

$$C^{CF}_o = \tilde{C}^{CF}_o e^{\frac{\Delta^1_o}{\sigma}}\ \forall\ o \in [2, \ldots, O] \tag{27}$$

where Δ^1_o is as defined in (24). For an arbitrary choice of g, we obtain:

$$P^{CF}(g|d(g), \tilde{\Theta}, \tilde{\mathbf{C}}^{\mathbf{CF}}) = \frac{e^{\frac{\tilde{\theta}^{CF}_g}{\sigma}} \overline{S}^{CF}_{g|o(g),d(g)} f^{CF}(o(g)) \tilde{C}^{CF}_o}{\sum_{o' \in \mathcal{O}} \sum_{g' \in (o',d)} e^{\frac{\tilde{\theta}^{CF}_{g'}}{\sigma}} \overline{S}^{CF}_{g'|o'(g'),d(g)} f^{CF}(o') \tilde{C}^{CF}_{o'}}$$

[51]Note that I have suppressed the dependence of $P^{CF}(*|\Theta, \mathbf{C}^{\mathbf{CF}}, f^{CF}(o), h^{CF}(d), \overline{S}_{g|o,d})$ on $f^{CF}(o)$, $h^{CF}(d)$, and $\overline{S}_{g|o,d}$ because these are held fixed across the two alternative counterfactual simulations.

$$= \frac{e^{\frac{(\theta_g^{CF}+\Delta^1_{o(g)}+\Delta^2_{d(g)})}{\sigma}}\overline{S}^{CF}_{g|o(g),d(g)}f^{CF}(o(g))C_o^{CF}e^{\frac{-\Delta^1_o}{\sigma}}}{\sum_{o'\in\mathcal{O}}\sum_{g'\in(o',d)}e^{\frac{(\theta_{g'}^{CF}+\Delta^1_{o(g')}+\Delta^2_{d(g')})}{\sigma}}\overline{S}^{CF}_{g'|o'(g'),d(g)}f^{CF}(o')C_{o'}^{CF}e^{\frac{-\Delta^1_{o'}}{\sigma}}}$$

$$= e^{\frac{\Delta^1_{o(g)}}{\sigma}}e^{\frac{\Delta^2_{d(g)}}{\sigma}}e^{\frac{-\Delta^1_{o(g)}}{\sigma}}\frac{e^{\frac{\theta_g^{CF}}{\sigma}}\overline{S}^{CF}_{g|o(g),d(g)}f^{CF}(o(g))C_o^{CF}}{e^{\frac{\Delta^2_{d(g)}}{\sigma}}\sum_{o'\in\mathcal{O}}e^{\frac{\Delta^1_{o(g')}}{\sigma}}e^{\frac{-\Delta^1_{o(g')}}{\sigma}}\sum_{g'\in(o',d)}e^{\frac{\theta_{g'}^{CF}}{\sigma}}\overline{S}^{CF}_{g'|o'(g'),d(g)}f^{CF}(o')C_{o'}^{CF}}$$

$$= \frac{e^{\frac{\theta_g^{CF}}{\sigma}}\overline{S}^{CF}_{g|o(g),d(g)}f^{CF}(o(g))C_o^{CF}}{\sum_{o'\in\mathcal{O}}\sum_{g'\in(o',d)}e^{\frac{\theta_{g'}^{CF}}{\sigma}}\overline{S}^{CF}_{g'|o'(g'),d(g)}f^{CF}(o')C_{o'}^{CF}} = P^{CF}(g|d,\Theta,\mathbf{C^{CF}}) \tag{28}$$

This proves that $P^{CF}(g|d,\Theta,\mathbf{C^{CF}})$ also satisfies the market clearing conditions (26) 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 Θ^{D-in-D} will in fact equal the counterfactual assignment I desire, which is based on the true set of joint surplus values Θ. Furthermore, while origin-type specific mean utility values $\tilde{\mathbf{C}}^{\mathbf{CF}}$ that clear the market given $\tilde{\Theta}$ will differ for each origin type from the corresponding vector $\mathbf{C^{CF}}$ based on the true surplus set Θ, 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 σ

I attempt to estimate σ, the standard deviation of the unobserved match-level component $\epsilon_{ij(i)k}$, by exploiting the evolution in the composition of U.S. origin and destination job matches $f^y(o)$ and $h^y(d)$ across years y. Specifically, I estimate the set of group-level surpluses $\{\theta_g^{2002}\}$ from the observed 2002-2003 matching. Then, holding these surplus values fixed, I combine $\{\theta_g^{2002}\}$ with $f^y(o)$ and $h^y(d)$ from each other year $y \in [2003, 2010]$ to generate counterfactual assignments and changes in scaled mean (exponentiated) utility values $\{C_o^{CF}\}$ for each origin type. These counterfactuals predict how mean worker utilities by skill/location combination would have evolved given the observed compositional changes in labor supply and demand had the underlying surplus values $\{\theta_g\}$ been constant and equal to $\{\theta_g^{2002}\}$ 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 $C_o^{CF} \approx \frac{1}{|o|}\sum_{i:o(i,j(i))=o} e^{\frac{-r_i^{CF}}{\sigma}}$. 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 $\frac{1}{|o^y|}\sum_{i:o(i,j(i))=o} e^{\frac{-r_i^{CF,y}}{\sigma^y}} \approx e^{\frac{\overline{r}_o^{CF,y}}{\sigma^y}}$, then taking logs yields $ln(C_o^{CF,y}) \approx \frac{\overline{r}_o^{CF,y}}{\sigma^y}$.

Next, I form the corresponding changes in observed annual earnings from origin to destination match for each origin type in each year, $\overline{Earn}_o^{y+1} - \overline{Earn}_o^y$.[52] I then run the following regression at the o-type level for each year $y \in [1993-2011]$:

$$\overline{Earn}_o^{y+1} - \overline{Earn}_o^y = \beta_0^y + \beta_1^y(ln(C_o^{CF,y+1}) - ln(C_o^{CF,y})) + \nu_o^y \tag{29}$$

Recall that the $\overline{r}_o^{CF,y}$ 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 to jobs featuring better or worse amenities, avoided more moving/recruiting training costs, 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 $\overline{r}_o^{CF,p,y+1} - \overline{r}_o^{CF,p,y}$ should approximately equal $\overline{Earn}_o^{y+1} - \overline{Earn}_o^y$. This implies that $\beta_1^y \approx \sigma^y$.

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 equation (29) for the collapsed origin type spaces associated with each possible target PUMA among the sample states, and averaged the estimates of β_1 across all regressions satisfying a minimum R^2 threshold of .1 to obtain $\hat{\beta}_1^y$.[53] The estimates of $\hat{\beta}_1^y$ are fairly consistent across years, so I use the mean estimate across all years, $\overline{\sigma} = 19,600$, 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 σ. 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 θ_g to evolve in a particular parametric fashion.[54] In fact, GS discuss how a vector of σ values associated with different types or combinations of types based on observed char-

[52] 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.

[53] 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, and β_1^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.

[54] In the actual implementation, I do allow the set of θ_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 decline in job-to-job mobility chronicled by Hyatt et al. (2016) during this time period is strongly at odds with the assumption that θ_g is completely time invariant.

acteristics might potentially be jointly estimated with other model parameters (thereby allowing heteroskedasticity across types in the idiosyncratic match component). Since the focus in this paper 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 CS's assignment model, the unobserved match-level heterogeneity is assumed to take the form $\epsilon_{ik} = \epsilon^1_{o(i)k} + \epsilon^2_{id(k)}$, so that aggregate surplus is left unchanged when two pairs of job matches (i, k) 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 $w_{ik} = w_{g(i,k)} \; \forall \, (i, k)$.[55] GS show that under this assumption, observing the (common) group-level transfers w_g would be sufficient to decompose the group-level mean joint surplus θ_g into the worker and position's respective pre-transfer payoffs, which I denote θ^l_g and θ^f_g, respectively.

Because the model proposed in section 2.2 does not impose the additive separability assumption $\epsilon_{ik} = \epsilon^1_{o(i)k} + \epsilon^2_{id(k)}$, equilibrium transfers will in general vary among (i, k) pairs within the same group g. Given the substantial earnings variance within observed groups g regardless of the worker, position, and job transition characteristics used to define g, the CS restriction on the nature of unobserved match-level heterogeneity would be strongly rejected in the labor market context.

However, one can still consider whether the observed transfers $\{w_{ik}\}$ identify additional objects. From section 2.1, equilibrium transfers are related to equilibrium worker and position payoffs via:

$$w_{ik} = \pi^f_{ik} - q_k \tag{30}$$

$$w_{ik} = r_i - \pi^l_{ik} \tag{31}$$

Next, recall from equation (11) that under Assumptions 1 and 2 in Proposition A1 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:

$$\ln(\frac{P(g_1|d)}{P(g_2|d)}) = \ln(P(g_1|d)) - \ln(P(g_2|d)) = \frac{\theta_{g_1}}{\sigma} + \ln(\overline{S}_{g_1|o(g_1),d}) + \ln(f(o(g_1))) + \ln(C_{o(g_1)}) - \frac{\theta_{g_2}}{\sigma} - \ln \overline{S}_{g_2|o(g_2),d}) - \ln(f(o(g_2))) - \ln(C_{o(g_2)}) \tag{32}$$

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

[55] 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.

sample is taken, extremely precisely estimated), one can form adjusted log odds:

$$\ln(\frac{\hat{P}_{g_1|d}/(\overline{S}_{g_1|o(g_1),d})f(o(g_1)))}{\hat{P}_{g_2|d}/(\overline{S}_{g_2|o(g_2),d}f(o(g_2)))}) = (\frac{\theta_{g_1}-\theta_{g_2}}{\sigma}) + (\ln(C_{o(g_1)}) - \ln(C_{o(g_2)})) \tag{33}$$

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

$$C_o = \frac{1}{|o|}\sum_{i:o(i,j(i))=o(g)} e^{-\frac{r_i}{\sigma}} \approx \sum_{\frac{1}{g_k}}\sum_{i:g(i,j(i),k)=g} e^{-\frac{r_i}{\sigma}} \;\forall\, k \tag{34}$$

Plugging (31) into (34) and then (34) into (33) yields:

$$\begin{aligned}&\ln(\frac{\hat{P}_{g_1|d}/(\overline{S}_{g_1|o(g_1),d}f(o(g_1)))}{\hat{P}_{g_2|d}/(\overline{S}_{g_2|o(g_2),d}f(o(g_2)))})\\ &= (\frac{\theta_{g_1}-\theta_{g_2}}{\sigma}) + (\ln(\frac{1}{|o|}\sum_{i:o(i,j(i))=o(g_1)} e^{-\frac{w_{ik}+\pi^l_{ik}}{\sigma}}) - \ln(\frac{1}{|o|}\sum_{i:o(i,j(i))=o(g_2)} e^{-\frac{w_{ik}+\pi^l_{ik}}{\sigma}}))\end{aligned} \tag{35}$$

It is not immediately obvious how to use equation (37) 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 CS. Specifically, suppose the following assumptions hold:

$$\begin{aligned}&r_i \approx r_{o(i)} \;\forall\, i : o(i,j(i)) = o \;\forall o \in \mathcal{O}\\ &\pi^l_{ik} = \pi^l_{g(i,k)} \equiv \theta^l_g \;\forall\, (i,k) : g(i,k) = g \;\forall\, g \in \mathcal{G}\\ &w_{ik} = w_{g(i,k)} \forall\, (i,k) : g(i,k) = g \;\forall\, g \in \mathcal{G}\end{aligned} \tag{36}$$

These assumptions are extremely unlikely to hold in any stable matching if there is meaningful variance in ϵ_{ik} among the (i,k) pairs within the same group g. Nonetheless, they yield:

$$\begin{aligned}&\ln(\frac{\hat{P}_{g_1|d}/(\overline{S}_{g_1|o(g_1),d}f(o(g_1)))}{\hat{P}_{g_2|d}/(\overline{S}_{g_2|o(g_2),d}f(o(g_2)))}) = (\frac{\theta_{g_1}-\theta_{g_2}}{\sigma}) + (\ln(e^{-r_{o(g_1)}}) - \ln(e^{-r_{o(g_2)}}))\\ &= (\frac{\theta_{g_1}-\theta_{g_2}}{\sigma}) + \frac{-r_{o(g_1)}+r_{o(g_2)}}{\sigma} = (\frac{\theta_{g_1}-\theta_{g_2}}{\sigma}) + (\frac{-(w_{g_1}+\theta^l_{g_1}) + (w_{g_2}+\theta^l_{g_2})}{\sigma})\\ &= \frac{\theta^f_{g_1}-\theta^f_{g_2}+(w_{g_2}-w_{g_1})}{\sigma}\end{aligned} \tag{37}$$

Given an estimate of σ based on multiple markets (as described in Appendix A3) and data on mean annual earnings for each transition group $g \in \mathcal{G}$, 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 θ^f_g 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 $\theta^l_{g_1} - \theta^l_{g_2}$ 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. Again, 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 (36) to be antithetical to the spirit of the model and at odds with the data, and 2) other than estimating σ, 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 $(y-1, y)$.

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 workers who are not observed working in a given year in my sample (despite being observed in other years) are in fact working in an 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 the procedure 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 one used to impute unemployment-to-unemployment transitions (U-to-U), I discuss the two separately.

A5.1 U-to-E and E-to-U 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 ≥ 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.[56] I

[56]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

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. $P(\text{True U-to-E}|\text{loc(o), age(o), loc(d), industry(d)})$. I then multiply the LEHD count for each transition group associated with U-to-E transitions by the estimate $P(\text{True U-to-E}|\text{loc(o), age(o), loc(d), industry(d)})$ 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 "unemployment" 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-OLF 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 sum of the counts of true and "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 ($P(\text{True E-to-U}|\text{loc(o), earn(o), age(o)})$).

I then multiply the LEHD count for each transition group associated with E-to-U transitions by the estimate $P(\text{True E-to-U}|\text{loc(o), earn(o), age(o)})$ 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 mismeasured in the LEHD, I rely particularly heavily on ACS and BLS data to generate these counts. I first create 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

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.

and U-to-U transitions, I then rescale these counts by multiplying by (# U-to-U BLS)/(# NE-to-U ACS). Here, (# NE-to-U ACS) is the sum of NE-to-U transition counts across all state-age combos for the chosen year, and (# U-to-U BLS) is the average across the year's final three months of the BLS reported count of workers 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, so that the sum of these imputed counts will at least match the BLS long-term unemployment count in the chosen year.[57]

Finally, since some U-to-U transition groups used in the model feature 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 across tracts and PUMAs the U-to-U counts that were originally computed at the state-age level.

A6 Smoothing Procedure

In this appendix I describe how I smooth the empirical distribution of transitions across groups, $\hat{P}(g)$, prior to estimation in order to generate accurate estimates of the set of identified joint surplus difference-in-differences Θ^{D-in-D}. I smooth for two reasons. First, such smoothing serves as a "noise infusion" technique that removes the risk that individual or establishment identities could be revealed by any estimates presented in the paper, as required of all 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 such that many transition groups are rarely (or never) observed in a given matching 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 transition belongs to a particular transition group is a smooth function of the observed characteristics that define group g (following Hotz and Miller (1993) and Arcidiacono and Miller (2011)). 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,

[57] Because I use a 50% random sample of LEHD transitions in this version, I multiply estimated U-to-U counts by .5.

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 θ_g whose estimate I wish to make more precise.

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 unemployment 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 $z(i, k)$ for whether establishment $j(i)$ and establishment k are the same, so that worker i is a job stayer rather than a mover (denoted $stayer(g)$).

Given the goal of accurately characterizing incidence at a very high spatial resolution, 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 the kernel estimator should place non-zero weight only on alternative groups g' that share the same origin and destination locations ($loc(o(g)) = loc(o(g'))$ and $loc(d(g)) = loc(d(g'))$). Similarly, I posit that an establishment's combination of size, average worker pay, and industry is likely to be more important than its location in determining the skill category of worker 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:

$$\begin{aligned}
&P(g) = P(g|d(g))h(d(g)) = P([o(g), d(g), z(g)]|d)h(d(g))\\
&= P([loc(o(g)), earn(o(g)), stayer(g)]|d)h(d(g))\\
&= P(loc(o(g))|earn(o(g)), stayer(g), d)P([earn(o(g)), stayer(g)]|d)h(d(g))\\
&= 1(stayer(g) = 1)P(loc(o(g))|earn(o(g)), 1(stayer(g) = 1), d)P([earn(o(g)), 1(stayer(g) = 1)]|d)h(d(g))\\
&+ 1(stayer(g) = 0)P(loc(o(g))|earn(o(g)), 1(stayer(g) = 0), d)P([earn(o(g)), 1(stayer(g) = 0)]|d)h(d(g))\\
&= 1(stayer(g) = 1)1(loc(o(g)) = loc(d(g)))P([earn(o(g)), 1(stayer(g) = 1)]|d)h(d(g))\\
&+ 1(stayer(g) = 0)P(loc(o(g))|earn(o(g), 1(stayer(g) = 0), d)P([earn(o(g)), 1(stayer(g) = 0)]|d)h(d(g))
\end{aligned} \tag{38}$$

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 $z(g) \equiv stayer(g)$ only takes on two values (0 for job movers and 1 for stayers), and the last line uses the fact that $P(loc(o(g))|earn(o(g)), 1(stayer(g) = 1), d) = 1(loc(o(g)) = loc(d(g)))$, since a potential stayer associated with a particular destination type must have already been working at the same location in the origin period (since I treat establishments that switch locations as different establishments for computational reasons). I use separate kernel density estimator procedures to estimate each of $P(loc(o(g))|earn(o(g)), 1(stayer(g) = 0), d(g))$, $P(earn(o(g)), 1(stayer(g) = 0)|d(g))$, and $P(earn(o(g)), 1(stayer(g) = 1)|d(g))$.

Consider first the estimation of $P(loc(o(g))|earn(o(g), 1(stayer(g) = 0), d(g))$, the conditional probability that a particular new hire would be originally located at location $loc(o)$, given the hired worker's initial earnings category and the destination position's type d. Let $K^{dist}(g,g')$ denote the metric capturing how similar an alternative group g' is to g for the purpose of estimating the propensity for establishments 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 $K^{dist}(g,g') < \infty$ (corresponding to non-zero weight) to empirical conditional probabilities $P(loc(o(g'))|earn(o(g')), 1(stayer(g') = 0), d(g'))$ of alternative groups g' that feature both the same origin location $loc(o(g')) = loc(o(g))$ *and* destination location $loc(d(g')) = loc(d(g))$.[58]

$K^{dist}(g,g')$ 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 skill category of hired workers. The closer $earn(o(g'))$ is to $earn(o(g'))$, the smaller is the assigned distance $K^{dist}(g,g')$, but the profile flattens so that all groups g' that differ from g' only due to $earn(o(g'))$ contribute to the weighted average. $K^{dist}(g,g')$ assigns larger (but still noninfinite) distance to groups g' featuring destination types that also differ on establishment size, avg. pay, 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 $earn(o(g'))$, denoted $N^{dist}(g')$ below, since this determines the signal strength of the empirical CCP $P(loc(o(g')|earn(o(g')), 1(stayer(g') = 0), d(g'))$. Thus, we have:

$$P(loc(o(g))|earn(o(g)), 1(stayer(g) = 0), d(g)) \approx$$
$$\sum_{g'} \left(\frac{\phi(K^{dist}(g', g)N^{dist}(g'))}{\sum_{g''} \phi(K^{dist}(g'', g)N^{dist}(g''))} \hat{P}(loc(o(g'))|earn(o(g')), 1(stayer(g') = 0), d(g')) \right) \quad (39)$$

where $\phi(*)$ is the normal density function (used as the kernel density), and $\frac{\phi(K^{dist}(g',g)N^{dist}(g'))}{\sum_{g''} \phi(K^{dist}(g'',g)N^{dist}(g''))}$ represents the weight given to a particular nearby transition group g'.[59]

Next, consider the estimation of $P(earn(o(g)), 1(stayer(g) = 1)|d)$ and $P(earn(o(g)), 1(stayer(g) = 0)|d)$, the conditional probabilities that either a job stayer or mover originally paid at a particular earnings quartile (or possibly unemployed for movers) will be hired to fill a position of destination type d. Let $K^{earn/move}(g,g')$ and $K^{earn/stay}(g,g')$ 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.

$K^{earn/move}(g,g')$ and $K^{earn/stay}(g,g')$ each assign infinite distance (i.e. zero weight) to groups

[58]There are a very small number of destination and origin types that are never observed in any transition. By necessity, I put positive weight on groups featuring nearby origin or destination locations in such cases.

[59]A standard deviation of 10 was used as the bandwidth choice for both this and the kernel densities presented below. The results were insensitive to moderate changes in bandwidth choice, though choosing a very small bandwidth resulted in very volatile simulation estimates across target tracts, highlighting the need for smoothing.

g' featuring different combos of establishment size, average pay, and industry than the target group g. $K^{earn/move}(g, g')$ ($K^{earn/stay}(g, g')$) assigns small distances to the conditional probabilities for groups g' representing hiring new (retaining) workers from the same initial earnings (or nonemployment) category $earn(o(g)) = earn(o(g'))$ among firms from the same destination type $d(g) = d(g')$ but who are hiring nearby workers. The distance metric increases in the tract pathlength between $loc(o(g'))$ and $loc(o(g'))$, but flattens beyond a threshold distance, so that groups featuring all origin locations (but shared values of other characteristics) contribute to the estimate.

Larger (but finite) distance values for $K^{earn/move}(g, g')$ and $K^{earn/stay}(g, g')$) 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 and average earnings quartiles and industry supersector. Again, the distance metric increases in the pathlength between $loc(d(g))$ and $loc(d(g'))$, 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 retention/new hire decision (conditional on the utility bids required by workers in different locations) is likely to be driven more by an establishment's production process (proxied by size, mean pay, and industry) than by its location. Since there still may be spatially correlated unobserved heterogeneity in production processes conditional on the other establishment observables, I place greater weight on the skill/retention decisions of proximate firms. More distant firms receive non-negligible weight only when too few local observations exist to form reliable estimates. The estimators for $P(earn(o(g)), 1(stayer(g) = 1)|d)$ and $P(earn(o(g)), 1(stayer(g) = 0)|d)$ can be expressed via:

$$P(earn(o(g)), 1(stayer(g) = 0)|d(g)) \approx$$
$$\sum_{g'} \left(\frac{\phi(K^{earn/move}(g', g)h(d(g')))}{\sum_{g''} \phi(K^{earn/move}(g'', g)h(d(g'')))} \hat{P}(earn(o(g')), 1(stayer(g') = 0)|d(g')) \qquad (40)$$
$$P(earn(o(g)), 1(stayer(g) = 1)|d(g)) \approx$$
$$\sum_{g'} \left(\frac{\phi(K^{earn/stay}(g', g)h(d(g')))}{\sum_{g''} \phi(K^{earn/stay}(g'', g)h(d(g'')))} \hat{P}(earn(o(g')), 1(stayer(g') = 1)|d(g')) \qquad (41)$$

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 CCPs without imposing assumptions about the spatial links between locations. Second, at the same time, one can still use information contained in the hiring and retention choices of more distant establishments to learn about the propensity for establishments of different sizes, pay levels, and industries to retain and hire workers at different skill levels and from unemployment. 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 establishments associated with groups featuring very similar characteristics to yield reliable estimates. Fourth, overall the estimated probabilities $P(g|d)$ place weight on many groups, so that no element of the resulting smoothed distribution contains identifying worker or establishment information, eliminating disclosure risk.

A7 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 job transitions driven by small firm expansions/contractions and worker retirements and preference or skill changes over the life cycle that generate considerable offsetting churn in the U.S. labor market. Thus, one might reasonably wonder whether parameters governing ordinary worker flows are capable of capturing the response to sizable, locally focused positive or negative shocks. To address this concern, in this section I describe and present results from a model validation exercise in which surplus parameters estimated on pre-shock ordinary worker flows were used to forecast the reallocation of workers after actual local economic shocks observed in the LEHD sample.

Specifically, 514 shocks to employment in a census tract 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; and 6) less than 50% of the change in number of positions filled in the year of the shock was 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 any worker reallocation, that the shock was big enough to represent a meaningful 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.

I assess the accuracy of the forecast using 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 sums 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 and multiplies by one-half: $\sum_g \frac{1}{2}|\hat{P(g)} - P(g)|$.

To help understand the sources of improvements and shortfalls in model fit, I also compute the index of dissimilarity between the true allocation and three 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 $P^y(g|d) = \frac{e^{X_g^y \lambda}}{\sum_{g'} e^{X_{g'}^y \lambda}}$, where X_g^y includes a substantial set of regressors constructed for year y that capture the kinds of predictors of joint surplus that researchers often use, and λ is the corresponding vector of parameters estimated from the relationship between the previous year's data, $P^{y-1}(g|d)$ and X_g^{y-1}. 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, and 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 ($1(z(g) = 1)$), the origin type frequency $f(o(g))$ interacted with the geographic category of the destination type associated with g (tract, PUMA, or state), an interaction between $f(o(g))$ and an indicator for whether $d(g)$ represents the "nonemployment" position type, and dummies for whether the origin and destination types associated with transition group g share a PUMA and share a state.

The second alternative forecast simply imposes that the CCPs that existed between $y-2$ and $y-1$ also hold during the shock year, so that $P^y(g) = \hat{P}^{y-1}(g|d)h^y(d)$. The third alternative forecast mimics the second, except that the smoothing procedure described in Section A6 is applied to the $y-2$ data prior to constructing $\hat{P}^{y-1}(g|d)$. Like much research on either worker job search or firm job filling, all these alternative forecasts ignore the problem's two-sided nature, and thus do not impose that the proposed allocation satisfies the marginal distribution of origin types, $f^{y-1}(o)$.

Table A20 contains the results of this exercise. All entries consist of averages across the 514 shocks considered. The first five columns in Row 1 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 from 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 originally in the relevant PUMA were misallocated. However, predicting the exact joint distribution of origin tract and initial skill category among workers 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 must be reallocated to a different group to match the actual post-shock allocation. Holding fixed the full prior year CCP distribution (cols. 3 and 8) performs slightly worse than the two-sided estimator within the target PUMA (37.5% misallocated), while smoothing the CCPs 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 are combined that feature the same worker and establishment characteristics as well as origin and destination locations that belong to the same distance bin (using 42 bins), so that the dissimilarity index is computed over a somewhat coarser set of 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 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 better matches 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 (row 4) provides a slight improvement, 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. 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.

Online Appendix Tables

Table A1: 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)

Distance from Target Tract	Share of JtJ Dest.	Initial Locations	Prob. of Stim. Job	Share of Stim Jobs	Change in P(Employed)	Share of Emp. Gains	Avg. Welfare Change ($)	Share of Wel. Gains
Target Tract	0.035	4.5E-05	0.015	0.031	0.003	0.005	1045	0.005
			(1.1E-04)	(9.2E-05)	(1.7E-05)	(1.4E-05)	(25)	(1.9E-05)
1 Tct Away	0.076	2.6E-04	0.005	0.057	9.4E-04	0.012	395	0.010
			(3.3E-05)	(1.5E-04)	(4.7E-06)	(2.6E-05)	(7)	(3.5E-05)
2 Tcts Away	0.072	5.9E-04	0.002	0.067	6.0E-04	0.017	278	0.014
			(1.1E-05)	(1.4E-04)	(2.2E-06)	(3.2E-05)	(3)	(3.9E-05)
3+ Tcts w/in PUMA	0.138	0.003	0.001	0.138	3.8E-04	0.047	164	0.041
			(6.7E-06)	(1.8E-04)	(1.3E-06)	(6.1E-05)	(1)	(7.9E-05)
1 PUMA Away	0.088	0.004	5.5E-04	0.096	2.6E-04	0.046	164	0.042
			(1.2E-06)	(1.6E-04)	(4.4E-07)	(6.9E-05)	(0.7)	(8.5E-05)
2 PUMAs Away	0.147	0.011	2.8E-04	0.149	1.8E-04	0.095	143	0.083
			(4.5E-07)	(1.8E-04)	(2.7E-07)	(1.1E-04)	(0.4)	(1.3E-04)
3+ PUMAs w/in State	0.365	0.162	5.3E-05	0.355	6.1E-05	0.433	109	0.390
			(2.2E-07)	(4.1E-04)	(1.7E-07)	(4.7E-04)	(0.5)	(5.0E-04)
1 State Away	0.031	0.135	8.3E-06	0.053	2.0E-05	0.128	89	0.132
			(2.0E-08)	(1.3E-04)	(2.0E-08)	(2.0E-04)	(0.0)	(2.1E-04)
2+ States Away	0.049	0.684	1.6E-06	0.053	6.7E-06	0.217	85	0.284
			(2.4E-09)	(7.9E-05)	(4.1E-09)	(1.8E-04)	(0.0)	(4.6E-04)

Notes: The column labeled “Share of JtJ Dest.” displays the share of all job-to-job transitions among 2010 and 2011 dominant jobs whose origin-destination distance fell into the distance bins given by the row labels. The column labeled “Initial Locations” captures 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). 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. 12/3+ Tct(s) Away” indicates that the origin establishment was one, two, or 3 or more tracts away (by tract pathlength) within the same PUMA. “1/2/3+ 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 A2: 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)

Distance from Centroid of Target Tract	Avg. Welfare Change ($)	Share of Wel. Gains	Change in P(Employed)	Share of Emp. Gains	Prob. of Stim. Job	Share of Stim. Jobs
Within 1 Mile	480	0.009	8.0E-04	0.008	0.004	0.036
1-2 Miles Away	230	0.014	4.4E-04	0.011	0.002	0.046
3-5 Miles Away	210	0.041	2.9E-04	0.037	9.3E-04	0.117
6-11 Miles Away	226	0.052	1.9E-04	0.047	5.1E-04	0.124
11-26 Miles Away	266	0.087	1.7E-04	0.077	3.7E-04	0.167
26-50 Miles Away	152	0.044	1.6E-04	0.041	2.8E-04	0.072
51-100 Miles Away	107	0.055	1.6E-04	0.055	1.7E-04	0.061
101-250 Miles Away	82	0.165	5.4E-05	0.159	3.5E-05	0.103
>250 Miles Away	66	0.533	1.3E-05	0.565	6.4E-06	0.275

Notes: See Table A1 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 A3: 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)

Earnings	Share of Employment Gains					Share of Welfare Gains				
Category	Avg.	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi	Avg.	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi
UE $\leq$ Age 25	0.089	0.087	0.097	0.082	0.091	0.046	0.045	0.050	0.041	0.046
UE $>$ Age 25	0.322	0.315	0.329	0.315	0.328	0.057	0.057	0.060	0.054	0.058
1st Quartile	0.261	0.269	0.265	0.257	0.254	0.201	0.207	0.210	0.192	0.196
2nd Quartile	0.154	0.159	0.152	0.155	0.151	0.214	0.220	0.217	0.208	0.209
3rd Quartile	0.096	0.096	0.090	0.101	0.096	0.223	0.224	0.218	0.225	0.223
4th Quartile	0.076	0.071	0.068	0.086	0.079	0.260	0.248	0.245	0.280	0.268

Notes: See Table A10 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 $\leq$ 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 A4: 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)

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

Notes: See Table A1 for expanded definitions of the row labels. See Table A3 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 A5: 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)

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

Notes: See Table A1 for expanded definitions of the row labels. See Table A3 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 A6: 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)

Earnings		Industry							
Quintile	Avg.	Info.	Manu.	R/W Trd.	Oth. Serv.	Ed./Hlth	Lei/Hosp.	Gov.	Const.
UE $\leq$ Age 25	620	515	655	632	757	561	823	551	465
UE $>$ Age 25	1165	952	1090	1097	1413	1102	1374	1000	1294
1st Quartile	999	982	993	876	1282	968	1048	982	866
2nd Quartile	1063	1015	1097	958	1407	986	960	1110	971
3rd Quartile	1104	1062	1082	1017	1484	1001	969	1189	1026
4th Quartile	1242	1335	1283	1192	1640	1025	1127	1195	1139

Notes: See Table A3 for expanded definitions of distance bins captured by the row labels. See Table A8 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 A7: 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)

Earnings	Firm Size/Pay Level Combination			
Quintile	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi
NE $\leq$ Age 25	625	802	413	640
NE $>$ Age 25	1169	1297	1015	1181
1st Quartile	1243	1130	825	800
2nd Quartile	1341	1102	934	874
3rd Quartile	1272	997	1134	1012
4th Quartile	1081	872	1728	1287

Notes: See Table A3 for expanded definitions of employment status/earnings quartile categories captured by the row labels. See Table A10 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 A8: 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)

Distance from Focal Tract	Avg.	Industry: Info.	Manu.	Trd./Tns.	Oth. Serv.	Ed./Hlth	Lei/Hosp.	Gov.	Const.
Target Tract	0.003	0.002	0.002	0.002	0.003	0.002	0.003	0.002	0.003
	(1.7E-05)	(1.8E-05)	(1.9E-05)	(2.1E-05)	(2.2E-05)	(1.8E-05)	(1.9E-05)	(1.8E-05)	(1.9E-05)
1 Tct Away	9.4E-04	9.4E-04	9.8E-04	8.6E-04	0.001	9.5E-04	9.6E-04	9.1E-04	9.2E-04
	(4.7E-06)	(5.0E-06)	(5.1E-06)	(5.2E-06)	(6.2E-06)	(5.0E-06)	(5.5E-06)	(4.5E-06)	(6.5E-06)
2 Tcts Away	6.0E-04	5.8E-04	6.2E-04	5.4E-04	6.7E-04	6.4E-04	6.0E-04	5.9E-04	6.0E-04
	(2.2E-06)	(2.2E-06)	(2.5E-06)	(2.6E-06)	(2.7E-06)	(2.5E-06)	(2.7E-06)	(2.3E-06)	(2.6E-06)
3+ Tcts w/in PUMA	3.8E-04	3.6E-04	3.6E-04	3.5E-04	3.9E-04	4.3E-04	3.9E-04	3.7E-04	3.7E-04
	(1.3E-06)	(1.3E-06)	(1.3E-06)	(1.7E-06)	(1.4E-06)	(1.6E-06)	(1.8E-06)	(1.3E-06)	(1.3E-06)
1 PUMA Away	2.6E-04	2.5E-04	2.6E-04	2.5E-04	2.8E-04	2.7E-04	2.6E-04	2.6E-04	2.6E-04
	(4.4E-07)	(4.3E-07)	(4.8E-07)	(4.8E-07)	(5.2E-07)	(6.3E-07)	(5.0E-07)	(4.6E-07)	(6.3E-07)
2 PUMAs Away	1.8E-04	1.8E-04	1.8E-04	1.8E-04	1.8E-04	1.8E-04	1.7E-04	1.8E-04	1.8E-04
	(2.7E-07)	(2.7E-07)	(2.9E-07)	(3.0E-07)	(3.1E-07)	(2.8E-07)	(3.0E-07)	(3.0E-07)	(3.2E-07)
3+ PUMAs w/in State	6.1E-05	6.2E-05	6.1E-05	6.2E-05	6.1E-05	6.0E-05	6.0E-05	6.2E-05	6.1E-05
	(1.7E-07)	(1.8E-07)	(1.8E-07)	(1.7E-07)	(2.0E-07)	(1.7E-07)	(1.7E-07)	(1.9E-07)	(1.7E-07)
1 State Away	2.0E-05	2.0E-05	2.1E-05	2.0E-05	1.9E-05	2.0E-05	2.0E-05	2.0E-05	2.1E-05
	(2.0E-08)	(2.2E-08)	(2.4E-08)	(2.1E-08)	(2.1E-08)	(2.6E-08)	(2.2E-08)	(2.2E-08)	(2.5E-08)
2+ States Away	6.7E-06	6.7E-06	6.8E-06	6.8E-06	6.6E-06	6.7E-06	7.0E-06	6.6E-06	6.7E-06
	(4.1E-09)	(4.7E-09)	(5.0E-09)	(5.3E-09)	(4.9E-09)	(5.0E-09)	(5.1E-09)	(4.2E-09)	(5.4E-09)

Notes: See Table A1 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 representing different industry supersectors. 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. 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 A9: Share of Additional Employment Produced by Stimulus Among Geographic Areas Defined by Distances from the Focal Tract:
Stimuli Consist of 500 New Jobs at Firms in Different Industries (Averaged Across Firm Size/Firm Average Earnings Combinations)

Distance from		Industry							
Focal Tract	Avg.	Info.	Manu.	R/W Trd.	Oth. Serv.	Ed./Hlth	Lei/Hosp.	Gov.	Const.
Target Tract	0.005	0.005	0.005	0.005	0.007	0.005	0.005	0.005	0.005
	(1.4E-05)	(1.4E-05)	(1.6E-05)	(1.6E-05)	(2.0E-05)	(1.5E-05)	(1.5E-05)	(1.5E-05)	(1.8E-05)
1 Tct Away	0.012	0.012	0.012	0.011	0.012	0.012	0.012	0.011	0.011
	(2.6E-05)	(2.8E-05)	(3.0E-05)	(2.7E-05)	(3.4E-05)	(3.0E-05)	(3.0E-05)	(2.6E-05)	(2.7E-05)
2 Tcts Away	0.017	0.016	0.017	0.015	0.018	0.018	0.017	0.016	0.016
	(3.2E-05)	(3.3E-05)	(3.7E-05)	(3.2E-05)	(4.0E-05)	(4.0E-05)	(3.8E-05)	(3.4E-05)	(3.7E-05)
3+ Tcts w/in PUMA	0.047	0.047	0.047	0.043	0.051	0.050	0.048	0.047	0.047
	(6.1E-05)	(6.5E-05)	(6.7E-05)	(6.5E-05)	(7.1E-05)	(7.0E-05)	(6.7E-05)	(6.5E-05)	(7.1E-05)
1 PUMA Away	0.046	0.044	0.045	0.044	0.050	0.048	0.045	0.045	0.046
	(6.9E-05)	(7.2E-05)	(7.2E-05)	(7.5E-05)	(7.9E-05)	(7.7E-05)	(7.7E-05)	(7.1E-05)	(8.0E-05)
2 PUMAs Away	0.095	0.094	0.094	0.095	0.099	0.095	0.093	0.095	0.096
	(1.1E-04)	(1.1E-04)	(1.1E-04)	(1.2E-04)	(1.2E-04)	(1.1E-04)	(1.1E-04)	(1.2E-04)	(1.2E-04)
3+ PUMAs w/in State	0.433	0.438	0.431	0.439	0.429	0.427	0.428	0.440	0.430
	(4.7E-04)	(4.9E-04)	(4.8E-04)	(4.9E-04)	(4.7E-04)	(4.8E-04)	(4.8E-04)	(4.8E-04)	(4.9E-04)
1 State Away	0.128	0.128	0.131	0.129	0.123	0.130	0.127	0.127	0.131
	(2.0E-04)	(2.1E-04)	(2.1E-04)	(2.1E-04)	(2.0E-04)	(2.1E-04)	(1.9E-04)	(2.0E-04)	(2.2E-04)
2+ States Away	0.217	0.216	0.218	0.220	0.212	0.216	0.225	0.213	0.216
	(1.8E-04)	(2.0E-04)	(2.1E-04)	(2.1E-04)	(2.0E-04)	(2.0E-04)	(2.1E-04)	(1.9E-04)	(2.2E-04)

Notes: See Table A1 for expanded definitions of the row labels. See Table A8 for expanded definitions of industry supersectors listed in column labels.

Table A10: 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)

Distance from	Change in P(Employed)				Share of Emp. Gains			
Focal Tract	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi
Target Tract	0.003	0.003	0.003	0.003	0.007	0.006	0.005	0.006
1 Tct Away	0.001	0.001	0.001	0.001	0.012	0.013	0.011	0.012
2 Tcts Away	6.2E-04	6.9E-04	5.4E-04	6.3E-04	0.017	0.019	0.015	0.017
3+ Tcts w/in PUMA	3.4E-04	3.9E-04	3.1E-04	3.7E-04	0.046	0.053	0.042	0.049
1 PUMA Away	2.5E-04	2.8E-04	2.4E-04	2.7E-04	0.045	0.049	0.042	0.047
2 PUMAs Away	1.7E-04	1.9E-04	1.6E-04	1.9E-04	0.090	0.101	0.087	0.100
3+ PUMAs w/in State	5.4E-05	5.9E-05	5.4E-05	5.9E-05	0.414	0.447	0.415	0.451
1 State Away	2.2E-05	1.8E-05	2.2E-05	1.8E-05	0.137	0.116	0.141	0.117
2+ States Away	7.1E-06	6.1E-06	7.4E-06	6.2E-06	0.229	0.196	0.238	0.199

Notes: See Table A1 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 supersectors (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 A11: 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)

Distance from	Avg. Welfare Change ($)				Share of Welfare Gains			
Focal Tract	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi	Sm./Low	Lg./Low	Sm./Hi	Lg./Hi
Target Tract	1197	1032	1122	998	0.005	0.004	0.004	0.004
1 Tct Away	400	388	418	400	0.010	0.010	0.010	0.010
2 Tcts Away	279	282	277	283	0.014	0.014	0.014	0.014
3+ Tcts w/in PUMA	188	196	188	196	0.039	0.042	0.039	0.043
1 PUMA Away	174	184	167	174	0.041	0.043	0.040	0.044
2 PUMAs Away	142	147	142	144	0.077	0.085	0.081	0.090
3+ PUMAs w/in State	122	126	121	122	0.373	0.403	0.376	0.409
1 State Away	91	90	90	84	0.141	0.123	0.142	0.120
2+ States Away	87	86	85	81	0.300	0.275	0.295	0.266

Notes: See Table A1 for expanded definitions of the row labels. See Table A10 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 A12: 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

Distance from	Change in P(Employed)					Share of Employment Gains				
Focal Tract	All	Rural	Urban	Small	Large	All	Rural	Urban	Small	Large
Target Tract	0.003	0.005	0.002	0.009	0.001	0.005	0.008	0.004	0.003	0.007
1 Tct Away	9.4E-04	0.002	3.9E-04	0.001	6.4E-04	0.012	0.017	0.006	0.011	0.013
2 Tcts Away	6.0E-04	0.001	3.3E-04	7.5E-04	5.1E-04	0.017	0.020	0.011	0.018	0.018
3+ Tcts w/in PUMA	3.5E-04	5.2E-04	2.1E-04	3.5E-04	3.4E-04	0.047	0.050	0.035	0.046	0.048
1 PUMA	2.6E-04	3.6E-04	1.6E-04	2.6E-04	2.8E-04	0.046	0.053	0.036	0.043	0.044
2 PUMAs Away	1.8E-04	2.5E-04	1.1E-04	1.7E-04	1.9E-04	0.095	0.113	0.072	0.092	0.095
3+ PUMAs w/in State	5.7E-05	6.7E-05	5.1E-05	5.3E-05	6.0E-05	0.433	0.284	0.610	0.467	0.411
1 State Away	2.0E-05	2.4E-05	1.5E-05	2.0E-05	2.1E-05	0.128	0.171	0.066	0.116	0.144
2+ States Away	6.7E-06	8.0E-06	5.5E-06	6.5E-06	6.8E-06	0.217	0.284	0.161	0.205	0.219

Notes: See Table A1 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 A13: 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

Distance from	Avg. Welfare Gain ($)					Share of Welfare Gains				
Focal Tract	All	Rural	Urban	Small	Large	All	Rural	Urban	Small	Large
Target Tract	1045	1724	799	1790	434	0.005	0.008	0.003	0.003	0.007
1 Tct Away	395	639	175	425	338	0.010	0.016	0.005	0.011	0.012
2 Tcts Away	278	431	155	247	289	0.014	0.017	0.008	0.017	0.015
3+ Tcts w/in PUMA	188	250	120	156	203	0.041	0.045	0.027	0.041	0.042
1 PUMA	164	207	110	138	180	0.042	0.051	0.029	0.036	0.044
2 PUMAs Away	143	182	103	121	158	0.083	0.101	0.059	0.085	0.084
3+ PUMAs w/in State	107	116	84	94	123	0.390	0.258	0.533	0.437	0.371
1 State Away	89	101	73	71	100	0.132	0.173	0.067	0.117	0.153
2+ States Away	85	95	71	67	96	0.284	0.332	0.267	0.254	0.271

Notes: See Table A1 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 A14: 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)

Distance from Target Tract	Change in P(Employed)		Share of Emp. Gains		Avg. Welfare Change ($)		Share of Wel. Gains	
	Unres.	Res.	Unres.	Res.	Unres.	Res.	Unres.	Res.
Target Tract	3.0E-03	1.4E-02	0.006	0.029	995	6938	0.004	0.020
1 Tct Away	1.2E-03	4.0E-03	0.015	0.049	399	1238	0.010	0.035
2 Tcts Away	7.5E-04	2.1E-03	0.021	0.059	297	759	0.015	0.045
3+ Tcts w/in PUMA	4.0E-04	7.3E-04	0.053	0.098	204	381	0.043	0.075
1 PUMA Away	2.8E-04	2.0E-04	0.050	0.035	189	383	0.041	0.030
2 PUMAs Away	1.9E-04	1.4E-04	0.102	0.074	152	209	0.082	0.059
3+ PUMAs w/in State	5.8E-05	4.5E-05	0.443	0.346	132	194	0.401	0.312
1 State Away	1.8E-05	1.6E-05	0.117	0.103	96	161	0.125	0.112
2+ States Away	6.0E-06	6.5E-06	0.194	0.210	93	158	0.277	0.312

Notes: See Table A1 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 A15: 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)

Distance from	Change in P(Unemployed)			Share of Emp. Loss			Change in Welfare ($)			Share of Wel. Loss		
Focal Tract	25%	50%	100%	25%	50%	100%	25%	50%	100%	25%	50%	100%
Target Tract	0.029	0.073	0.192	0.126	0.158	0.202	-5622	-10474	-17028	0.100	0.097	0.083
1 Tct Away	4.4E-04	8.1E-04	1.4E-03	0.011	0.010	0.009	-235	-361	-543	0.009	0.009	0.009
2 Tcts Away	2.5E-04	4.7E-04	8.0E-04	0.015	0.013	0.011	-192	-279	-385	0.012	0.013	0.012
3+ Tcts w/in PUMA	1.5E-04	2.8E-04	5.0E-04	0.042	0.039	0.034	-168	-231	-302	0.034	0.037	0.040
1 PUMA Away	1.0E-04	2.0E-04	3.8E-04	0.037	0.036	0.033	-153	-206	-271	0.031	0.033	0.037
2 PUMAs Away	6.6E-05	1.3E-04	2.6E-04	0.073	0.073	0.069	-141	-183	-228	0.061	0.067	0.075
3+ PUMAs w/in State	2.4E-05	4.8E-05	9.3E-05	0.386	0.377	0.357	-134	-168	-198	0.319	0.331	0.343
1 State Away	8.4E-06	1.7E-05	3.2E-05	0.111	0.108	0.103	-118	-137	-138	0.115	0.115	0.113
2+ States Away	3.0E-06	5.6E-06	1.1E-05	0.199	0.186	0.183	-117	-134	-133	0.319	0.299	0.289

Notes: See Table A1 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 A16: 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)

Earnings	Share of Emp. Loss			Share of Wel. Loss		
Quintile	25%	50%	100%	25%	50%	100%
UE $\leq$ Age 25	0.081	0.077	0.072	0.042	0.042	0.042
UE $>$ Age 25	0.267	0.254	0.237	0.053	0.051	0.050
1st Quartile	0.278	0.274	0.268	0.210	0.208	0.207
2nd Quartile	0.172	0.176	0.181	0.218	0.218	0.218
3rd Quartile	0.112	0.120	0.132	0.223	0.225	0.225
4th Quartile	0.089	0.098	0.111	0.254	0.256	0.258

Notes: See Table A3 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 A17: 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	25% of Jobs Destroyed						100% of Jobs Destroyed					
Focal Tract	UE $\leq$ 25	UE $>$ 25	1st Q.	2nd Q.	3rd Q.	4th Q.	UE $\leq$ 25	UE $>$ 25	1st Q.	2nd Q.	3rd Q.	4th Q.
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	4.4E-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 A1 for expanded definitions of the distance bins represented by the row labels. See Table A3 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 A18: 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

Distance from	25% of Jobs Destroyed						100% of Jobs Destroyed					
Focal Tract	UE $\leq$ 25	UE > 25	1st Q.	2nd Q.	3rd Q.	4th Q.	UE $\leq$ 25	UE > 25	1st Q.	2nd Q.	3rd Q.	4th Q.
Target Tract	-291	-312	-4755	-6043	-7022	-7531	-589	-620	-13950	-18000	-21550	-23400
1 Tct Away	-215	-192	-227	-241	-247	-245	-417	-348	-522	-570	-613	-565
2 Tcts Away	-197	-174	-190	-193	-197	-194	-370	-309	-385	-392	-397	-398
3+ Tcts w/in PUMA	-166	-157	-167	-168	-170	-170	-299	-264	-294	-307	-312	-308
1 PUMA Away	-156	-145	-152	-155	-155	-153	-265	-230	-265	-273	-279	-281
2 PUMAs Away	-142	-134	-139	-141	-141	-143	-230	-196	-222	-229	-234	-237
3+ PUMAs w/in State	-135	-129	-134	-134	-135	-135	-199	-179	-197	-199	-201	-202
1 State Away	-118	-118	-118	-118	-118	-118	-139	-136	-138	-138	-138	-138
2+ States Away	-117	-117	-117	-117	-117	-117	-133	-132	-133	-133	-133	-133

Notes: See Table A1 for expanded definitions of the distance bins represented by the row labels. See Table A3 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 $ 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 A19: 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)

Distance from	25% of Jobs Destroyed							100% of Jobs Destroyed						
Focal Tract	Overall	UE ≤ 25	UE > 25	1st Q.	2nd Q.	3rd Q.	4th Q.	Overall	UE ≤ 25	UE > 25	1st Q.	2nd Q.	3rd Q.	4th Q.
Unemployment	0.029	0.002	0.006	0.044	0.036	0.028	0.022	0.192	0.006	0.016	0.218	0.233	0.225	0.194
Target Tract	-0.106	-0.010	-0.010	-0.137	-0.127	-0.111	-0.106	-0.705	-0.025	-0.023	-0.671	-0.792	-0.856	-0.876
1 Tct Away	0.005	0.000	0.000	0.006	0.006	0.005	0.004	0.030	0.001	0.001	0.028	0.037	0.040	0.033
2 Tcts Away	0.005	0.000	0.000	0.007	0.006	0.005	0.004	0.029	0.000	0.000	0.030	0.036	0.037	0.029
3+ Tcts w/in PUMA	0.010	0.001	0.000	0.013	0.013	0.011	0.009	0.063	0.001	0.000	0.061	0.076	0.082	0.070
1 PUMA Away	0.008	0.001	0.000	0.010	0.009	0.009	0.007	0.051	0.002	0.001	0.048	0.058	0.068	0.058
2 PUMAs Away	0.012	0.001	0.001	0.014	0.014	0.012	0.012	0.078	0.003	0.001	0.073	0.088	0.095	0.096
3+ PUMAs w/in State	0.032	0.003	0.002	0.038	0.036	0.033	0.039	0.219	0.009	0.004	0.188	0.227	0.252	0.324
1 State Away	0.003	0.001	0.001	0.003	0.003	0.004	0.004	0.022	0.002	0.001	0.014	0.022	0.030	0.033
2+ States Away	0.003	0.001	0.000	0.002	0.003	0.003	0.004	0.021	0.002	0.001	0.012	0.017	0.025	0.038

Notes: See Table A1 for expanded definitions of the distance bins represented by the row labels. See Table A3 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 A20: 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

	All U.S.				Target PUMA Only			
Level of Group Aggregation	Two-Sided Matching	Param. Logit	Raw CCP	Smoothed CCP	Two-Sided Matching	Param. Logit	Raw CCP	Smoothed CCP
Full Group Space	0.071	0.427	0.066	0.064	0.362	0.504	0.375	0.356
	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)	(0.004)	(0.003)	(0.002)
Sm. Dist. Bins	0.053	0.419	0.050	0.047	0.211	0.455	0.229	0.202
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)
Sm. Dist. Bins & No Firm Char.	0.009	0.195	0.022	0.017	0.053	0.325	0.078	0.066
	(0.000)	(0.003)	(0.001)	(0.000)	(0.001)	(0.004)	(0.002)	(0.001)
Lg. Dist. Bins & No Firm Char.	0.007	0.184	0.014	0.013	0.047	0.306	0.070	0.058
	(0.000)	(0.002)	(0.000)	(0.000)	(0.001)	(0.003)	(0.002)	(0.001)
Unemp. Only (All Loc.)	0.059	0.148	0.067	0.075	0.146	0.162	0.144	0.156
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Unemp. Only (Lg. Dist. Bins)	0.043	0.125	0.052	0.054	0.078	0.120	0.073	0.088
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)

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.