Economica

Economica (2022) **89**, S110–S136
doi:10.1111/ecca.12387

Suburbanization in the USA, 1970–2010

By STEPHEN J. REDDING

Princeton University, NBER and CEPR

Final version received 30 July 2021.

The second half of the 20th century saw large-scale suburbanization in the USA, with the median share of residents who work in the county where they live falling from 87% to 71% between 1970 and 2000. We introduce a new methodology for discriminating between the three leading explanations for this suburbanization (workplace attractiveness, residence attractiveness and bilateral commuting frictions). This methodology holds in the class of spatial models that are characterized by a structural gravity equation for commuting. We show that the increased openness of counties to commuting is explained mainly by reductions in bilateral commuting frictions, consistent with the expansion of the interstate highway network and the falling real cost of car ownership. We find that changes in workplace attractiveness and residence attractiveness are more important in explaining the observed shift in employment by workplace and employment by residence towards lower densities over time.

INTRODUCTION

One of the most striking features of the USA in the period since the Second World War was suburbanization: the dispersion of population from central cities and an increase in commuting distances to work. By the year 2019, the average one-way commute to work in the USA was 27.6 minutes, and a record 9.8% of commuters reported daily one-way commutes of at least 1 hour.[1] Workers make this substantial daily investment, to live and work in different locations, so as to balance their living costs and residential amenities with the employment and income opportunities at their place of employment. The resulting sprawl of suburban neighbourhoods has been blamed for a number of ills, including: environment degradation from the destruction of open countryside; urban blight including increased crime and a decline in quality of public schools in central cities; increased segregation by levels of human capital, race and ethnicity; and deteriorating city finances as more of the tax base moves beyond city boundaries.

Although this decentralization of both population and employment in the USA is a widely accepted feature of the post Second World War period, there remains considerable debate about the economic forces underlying it, and there is no commonly accepted theoretical framework for assessing the relative contributions of these different forces. Our main contribution is to develop a methodology for discriminating between the three leading explanations for this observed increase in commuting: (i) workplace forces, including a dispersion of manufacturing from central cities; (ii) residence forces, including increased demand for residential floor space and changes in local amenities/disamenities, such as crime and local public schools; (iii) bilateral commuting costs, including the expansion of the interstate highway network. Our methodology holds in an entire class of spatial models that are characterized by a structural gravity equation for commuting, in which bilateral commuting flows depend on bilateral commuting frictions, a workplace fixed effect and a residence fixed effect. The key idea underlying our approach is to use the observed changes in commuting flows and the structure of the gravity equation from this class of models to reveal the relative importance of these different explanations.

We implement this methodology using population census data on bilateral commuting flows between US counties from 1970 to 2000. We are thus able to characterize the change in the overall distribution of economic activity within the USA, including variation both within and between metro areas, and between metro areas and rural counties. We document three large-scale changes in observed patterns of commuting in the late 20th century USA. First, counties became substantially more open to commuting flows. Between 1970 and 2000, the median share of residents who work in the county where they live fell from 87% to 71%, and the fraction of counties with values for this share of less than 50% increased almost fourfold, from around 5% to about 18%.[2] Second, there was a dispersion of both employment and population from central cities, with the distributions of both workplace and residence employment shifting towards lower densities over this time period. Third, the shift in the distribution was larger for workforce employment than for residence employment, with the result that the distribution of workplace employment became substantially less spatially concentrated over time. We show that many of these observed changes are driven by variation within metro areas.

We show how our structural gravity approach can be used to consistently estimate measures of workplace attractiveness, residence attractiveness and bilateral commuting frictions. We are able to undertake this estimation without making assumptions about the values of model parameters, such as the dispersion of idiosyncratic preferences, because these parameters are already incorporated into our estimates of each of these components. We also estimate these components without taking a stand on whether they are exogenous or endogenous, because each of these components is separately identified by the log additive structure of the commuting gravity equation. We use these estimates to undertake counterfactuals in which we change one or more of these components and solve for the new spatial equilibrium distribution of economic activity. We use these counterfactuals to evaluate the relative importance of each component or mechanism in the model, as in the macroeconomics literatures on growth and business cycle accounting. In these counterfactuals, we again do not take a stand on whether these components are exogenous or endogenous. Regardless of what form the underlying causal forces take, our counterfactuals isolate the relative importance of each component or mechanism in the model, as in the macroeconomics literature on growth and business cycle accounting.

We show that changes in bilateral commuting frictions are the dominant force explaining the observed increase in counties' openness to commuting over time. Holding workplace and residence attractiveness constant at their 2000 values, and changing only bilateral commuting frictions to their 1970 values, the median share of residents who work in the county where they live rises from 71% to 85%, almost as large as the rise to 87% observed in the data. This pattern of results points towards the importance of reductions in bilateral travel costs, associated for example with the falling real costs of car ownership and the continued expansion of the interstate highway system over this time period.

We show that this decline in bilateral commuting frictions is less important than changes in workplace and residence attractiveness in explaining the observed shift in employment by workplace and employment by residence towards lower densities over time. Holding bilateral commuting frictions and residence attractiveness constant at their 2000 values, and changing only workplace attractiveness to its 1970 values, the number of counties with shares of employment by residence below the 10th percentile of the 2000 distribution falls from 307 to 193, larger than the fall to 204 observed in the data. Similarly, holding bilateral commuting frictions and workplace attractiveness constant at their 2000 values, and changing only residence attractiveness to its 1970 values, the number of counties with shares of employment by residence below the same threshold falls from 307 to 246, compared to the fall to 204 observed in the data. In contrast, holding residence and workplace attractiveness constant at their 2000 values, and changing only bilateral commuting frictions to their 1970 values, the number of counties with shares of residence employment below the same threshold actually increases to 487. These findings for workplace and residence attractiveness are consistent with a role for forces such as the relocation of manufacturing and changes in amenities in central cities in explaining the observed movement of economic activity towards lower densities over time during our sample period.

We focus on counties throughout our empirical analysis because of the availability of publicly available data on the matrix of bilateral commuting flows between counties over a long historical time period. Counties also have two other advantages from the point of view of our empirical analysis. First, county boundaries are relatively stable from 1970 until 2000, whereas the boundaries of smaller spatial units, such as census tracts, are subject to greater changes over time. Second, most counties have thousands of residents, and hence issues of granularity from the realized values of variables departing from their expected values as a result of sampling variation are much less of a concern at the county level than for smaller spatial units. Finally, to the extent that openness to commuting also increased at finer spatial scales within counties, our results are likely to be conservative in understating the overall increase in openness to commuting over time.

Our research is related to three separate lines of work that have proposed transport improvements, the decline of manufacturing in central cities, and changes in amenities and local public goods in central cities as potential explanations for the observed decentralization of economic activity.

A first line of research has used quasi-experimental sources of variation to provide empirical evidence on the impact of transport improvements, as reviewed in Redding and Turner (2015). One group of studies has used variation across cities and regions, including Chandra and Thompson (2000), Michaels (2008), Duranton and Turner (2011, 2012), Faber (2014), Duranton *et al.* (2014), Donaldson and Hornbeck (2016), Donaldson (2018) and Baum-Snow *et al.* (2020). A second group of studies has looked within cities, including McDonald and Osuji (1995), Gibbons and Machin (2005), Baum-Snow and Kahn (2005), Billings (2011), Brooks and Lutz (2018), and Heblich *et al.* (2020). Particularly relevant is a third group of studies that have examined the impact of interstate highway rays from central cities. A robust finding from this empirical literature is that these transport improvements lead to a decentralization of economic activity that is larger for population than employment, including Baum-Snow (2007, 2020) and Baum-Snow *et al.* (2017, 2020).

A second strand of research has emphasized the decline of manufacturing employment in central cities as a result of changes in production technology, transportation costs and international trade. Whereas in the late 19th century, central cities in the USA contained major concentrations of manufacturing activity, tradable and non-tradable services now account for the vast majority of employment. A number of studies have drawn attention to this transformation of urban areas and the spatial distribution of manufacturing employment, including Brezis and Krugman (1997), Glaeser and Kohlhase (2004), Stevens and Holmes (2004), Desmet and Rossi-Hansberg (2009), Michaels *et al.* (2019) and Hornbeck and Rotemberg (2021).

A third group of studies have pointed towards changes in amenities, crime and public schools as important drivers of the dispersion of population from central cities. Research on crime and public schools includes Benabou (1996), Benabou and Fernandez (1996), Glaeser and Sacerdote (1999), and Eckert and Kleineberg (2021). A number of studies highlight increased segregation in terms of race, ethnicity and human capital, including Boustan (2000), Boustan and Margo (2009), Fogli and Guerrieri (2019), and Chetty *et al.* (2020). As our bilateral commuting data aggregate all types of workers together, we are unable to explore segregation by worker characteristics, and focus instead on characterizing changes in the overall distribution of workers and residents across locations. More recently, several studies have pointed to a reversal of postwar trends of suburbanization, with gentrification and increases in population density in central cities, including Guerrieri *et al.* (2013), Couture and Handbury (2020), and Couture *et al.* (2020). This resurgence of central cities is, however, concentrated in recent years after the end of our sample period in 2000.

Besides these three main candidate explanations for suburbanization, our paper is related to three additional areas of research.

First, it contributes to research on the internal structure of cities and the separation of workplace and residence. Early theoretical research in this area assumed a monocentric city structure, including Alonso (1964), Mills (1967) and Muth (1969). Subsequent theoretical research modelled non-monocentric organizations of economic activity in linear cities or symmetric circular cities, as in Fujita and Ogawa (1982), Fujita and Krugman (1995), and Lucas and Rossi-Hansberg (2002). More recent work has developed quantitative urban models, which capture the rich asymmetric organizations of economic activity observed in the data, including Ahlfeldt *et al.* (2015), Allen *et al.* (2017), Monte *et al.* (2018), Tsivanidis (2018), Owens *et al.* (2020) and Miyauchi *et al.* (2021), as recently reviewed in Redding and Rossi-Hansberg (2017).

Second, our work connects with the wider theoretical and empirical literature on economic geography and agglomeration, including Henderson (1974), Krugman (1991), Krugman and Venables (1995), Fujita *et al.* (1999), Fujita and Thisse (2002), Davis and Weinstein (2002), Rossi-Hansberg (2005), Redding and Sturm (2008), Allen and Arkolakis (2014), Kline and Moretti (2014), Redding (2016), Caliendo *et al.* (2018), Dingel and Tintelnot (2020), and Kleinman *et al.* (2021), as reviewed in Duranton and Puga (2004), Rosenthal and Strange (2004), Moretti (2011), Combes and Gobillon (2015), and Redding (2021).

Finally, our findings are related to a historical literature on suburbanization in the USA and other countries, including Warner (1978), Hershberg (1981), Jackson (1987), Jacobs (1992), Fogelson (2003), Rae (2005), Angel and Lamson-Hall (2014), and Lee (2020). We focus on the substantial suburbanization that occurred in the last three decades of the 20th century, during which the automobile was the dominant mode of transport. But this suburbanization is a continuation of a longer-term trend from improvements in transport technology, dating back to the railway era of the 19th century, as analysed in Heblich *et al.* (2020).

The remainder of the paper is structured as follows. Section I develops our theoretical framework. Section II introduces our data. Section III presents reduced-form evidence. Section IV reports our quantitative findings. Section V summarizes our conclusions.

I. Theoretical Framework

We consider an economy that consists of a discrete set of locations indexed by $n, i \in N$, which will correspond to counties in our empirical application below. Time is discrete and is indexed by t. The economy as a whole is populated with an exogenous continuous measure of workers ($\bar{L}_t$), who are geographically mobile and endowed with one unit of labour that is supplied inelastically. Workers simultaneously choose their preferred residence n and workplace i given their idiosyncratic preference draws for amenities. With a continuous measure of workers, the law of large numbers applies, and the expected values of variables for a given residence and workplace equal the realized values. We allow locations to differ from one another in terms of their attractiveness for production and residence, as determined by productivity, amenities and the supply of floor space, and transport connections, where each of these location characteristics can evolve over time.

Residence–workplace choice

The preferences of a worker ω who lives in location n and works in location i at time t depend on residence attractiveness ($\mathscr{R}_{nt}$), workplace attractiveness ($\mathscr{W}_{it}$), bilateral commuting frictions (κ_{nit}), and an idiosyncratic amenity draw that is specific to each worker and residence–workplace pair ($b_{nit}(\omega)$):

$$U_{nit}(\omega) = \frac{b_{nit}(\omega)}{\kappa_{nit}} \mathscr{R}_{nt} \mathscr{W}_{it}.$$

This specification of the worker's choice of residence and workplace encompasses a large class of spatial models. We use residence attractiveness ($\mathscr{R}_{nt}$) to capture the determinants of the attractiveness of a location as a residence, including amenities such as scenic views, disamenities including crime, local public goods such as schools, and the cost of housing. We use workplace attractiveness ($\mathscr{W}_{it}$) to capture anything about a workplace that it makes it a more attractive place to work, including the wage, compensating differentials and access to surrounding consumption opportunities. We use bilateral commuting frictions (κ_{nit}) to capture any feature of a bilateral commute that makes it more or less costly for the worker, including public transit options, travel time and congestion. For some individual pairs of locations (e.g. Manhattan and Los Angeles), these bilateral commuting frictions may be sufficiently large that the measure of commuters is negligible. The idiosyncratic amenity ($b_{nit}(\omega)$) captures all of the idiosyncratic factors that can lead individual workers to live and work in different locations.

In general, residence attractiveness, workplace attractiveness and bilateral commuting frictions are all endogenous. For example, the price of housing is included in residence attractiveness and is determined by the market clearing condition for housing in each location. Similarly, the wage is included in workplace attractiveness and is the solution to the market clearing condition for labour in each location. Additionally, both residence and workplace attractiveness can be influenced by agglomeration economies in the density of surrounding economic activity. Finally, bilateral commuting frictions include congestion, which depends on the spatial equilibrium distribution of economic activity across all locations. Despite the endogeneity of each of these three terms, we show how they can be consistently estimated using only the observed data on bilateral commuting flows and the structural gravity equation in this class of models. As in the literature estimating firm and worker fixed effects using linked employee–employer datasets, each of these terms is separately identified through the log additive structure of the gravity equation. Nevertheless, when we undertake counterfactuals for changes in each of these terms below, it is important to keep in mind that they are endogenous, and hence that these counterfactuals reveal the relative importance of different mechanisms in the model for the observed variation in the data, and do not necessarily capture causal effects of exogenous changes in these terms.

As in a long line of research in transportation and spatial economics following McFadden (1974), we assume that worker idiosyncratic preferences for each residence–workplace pair are drawn independently each period from an extreme value distribution

$$G(b) = \exp(-b^{-\varepsilon}),$$

where we normalize the Fréchet scale parameter to 1 because it enters the model isomorphically to residence and workplace attractiveness; a smaller Fréchet shape parameter ε implies greater heterogeneity in idiosyncratic amenities, which in turn implies that worker location decisions are less sensitive to economic variables.

A first key implication of this extreme value specification for idiosyncratic preferences is that the unconditional probability that a worker chooses to commute from residence n to workplace i follows a gravity equation

$$\lambda_{nit} = \frac{L_{nit}}{\bar{L}_t} = \frac{(\mathscr{R}_{nt}\mathscr{W}_{it}/\kappa_{nit})^{\varepsilon}}{\sum_{r\in N}\sum_{\ell\in N}(\mathscr{R}_{rt}\mathscr{W}_{\ell t}/\kappa_{r\ell t})^{\varepsilon}}, \tag{1}$$

where L_{nit} is the measure of commuters from residence n to workplace i.

Therefore the probability of commuting between residence n and workplace i depends on the characteristics of that residence n, the attributes of that workplace i, and bilateral commuting costs ('bilateral resistance'). Furthermore, this probability also depends on the characteristics of all residences r, all workplaces ℓ, and all bilateral commuting costs ('multilateral resistance'). A large reduced-form literature finds that the gravity equation provides a good approximation to observed commuting flows, as reviewed, for example, in Fotheringham and O'Kelly (1989) and McDonald and McMillen (2010).

Summing across workplaces i in these bilateral commuting probabilities (1), we obtain what we term the 'residence probability', namely the probability that a worker lives in residence n

$$\lambda^R_{nt} = \frac{\sum_{\ell\in N} L_{n\ell t}}{\bar{L}_t} = \frac{R_{nt}}{\bar{L}_t} = \frac{\sum_{\ell\in N}(\mathscr{R}_{nt}\mathscr{W}_{\ell t}/\kappa_{n\ell t})^{\varepsilon}}{\sum_{r\in N}\sum_{\ell\in N}(\mathscr{R}_{rt}\mathscr{W}_{\ell t}/\kappa_{r\ell t})^{\varepsilon}}, \tag{2}$$

where R_{nt} denotes the measure of residents in location n at time t. Rewriting these residence probabilities, they depend on residence attractiveness ($\mathscr{R}_{nt}$) and a measure of residents' commuting market access (RMA_{nt}), given by

$$\lambda^R_{nt} = \frac{\mathscr{R}^{\varepsilon}_{nt}\, RMA^{\varepsilon}_{nt}}{\sum_{r\in N}\mathscr{R}^{\varepsilon}_{rt}\, RMA^{\varepsilon}_{rt}},$$

where RMA_{nt} summarizes access to surrounding employment opportunities and equals the weighted average of workplace attractiveness ($\mathscr{W}_{it}$) in each location, using the commuting frictions ($\kappa_{nit}^{-\varepsilon}$) as weights:

$$RMA_{nt} = \left(\sum_{\ell \in N} \left(\mathscr{W}_{\ell t} / \kappa_{n\ell t} \right)^{\varepsilon} \right)^{1/\varepsilon}.$$

Summing across residences n in the bilateral commuting probabilities (1), we obtain what we term the 'workplace probability', namely the probability that a worker is employed in workplace i:

$$(3) \qquad \lambda_{it}^{L} = \frac{\sum_{r \in N} L_{rit}}{\bar{L}_t} = \frac{L_{it}}{\bar{L}_t} = \frac{\sum_{r \in N} \left(\mathscr{R}_{rt} \mathscr{W}_{it} / \kappa_{rit} \right)^{\varepsilon}}{\sum_{r \in N} \sum_{\ell \in N} \left(\mathscr{R}_{rt} \mathscr{W}_{\ell t} / \kappa_{r\ell t} \right)^{\varepsilon}},$$

where L_{it} denotes the measure of workers in location i at time t. Rewriting these workplace probabilities, they depend on workplace attractiveness ($\mathscr{W}_{it}$) and a measure of workers' commuting market access (WMA_{it}), given by

$$\lambda_{it}^{L} = \frac{\mathscr{W}_{it}^{\varepsilon} WMA_{it}^{\varepsilon}}{\sum_{\ell \in N} \mathscr{W}_{\ell t}^{\varepsilon} WMA_{\ell t}^{\varepsilon}},$$

where WMA_{it} summarizes access to surrounding residential opportunities and equals the weighted average of residence attractiveness ($\mathscr{R}_{nt}$) in each location, using the commuting frictions ($\kappa_{nit}^{-\varepsilon}$) as weights:

$$WMA_{it} = \left(\sum_{r \in N} \left(\mathscr{R}_{rt} / \kappa_{rit} \right)^{\varepsilon} \right)^{1/\varepsilon}.$$

From the commuting probabilities (1), the residence probability (2) and the workplace probability (3), we can compute two measures of the openness of a location to commuting flows, one by residence, and the other by workplace. The residence-based measure is the share of residents who work where they live, which we call the 'conditional residence probability':

$$(4) \qquad \lambda_{nnt|n}^{R} = \frac{\lambda_{nnt}}{\lambda_{nt}^{R}} = \frac{\left(\mathscr{W}_{nt} / \kappa_{nnt} \right)^{\varepsilon}}{\sum_{\ell \in N} \left(\mathscr{W}_{\ell t} / \kappa_{n\ell t} \right)^{\varepsilon}} = \left(\frac{\mathscr{W}_{nt} / \kappa_{nnt}}{RMA_{nt}} \right)^{\varepsilon}.$$

This conditional residence probability does not depend on residence attractiveness ($\mathscr{R}_{nt}^{\varepsilon}$), because it conditions on living in a given residence. Therefore residence attractiveness is the same across all possible choices of workplaces conditional on living in that residence, and cancels from the numerator and denominator of this conditional probability.

The corresponding workplace-based measure is the share of workers who live where they work, which we call the 'conditional workplace probability':

$$\lambda_{nnt|n}^{L} = \frac{\lambda_{nnt}}{\lambda_{nt}^{L}} = \frac{\left(\mathscr{R}_{nt} / \kappa_{nnt} \right)^{\varepsilon}}{\sum_{\ell \in N} \left(\mathscr{R}_{rt} / \kappa_{rnt} \right)^{\varepsilon}} = \left(\frac{\mathscr{R}_{nt} / \kappa_{nnt}}{WMA_{nt}} \right)^{\varepsilon}.$$

This conditional workplace probability does not depend on workplace attractiveness ($\mathscr{W}_{it}^{\varepsilon}$), because it conditions on working in a given workplace. Therefore workplace attractiveness ($\mathscr{W}_{it}^{\varepsilon}$) is the same across all possible choices of residence conditional on working in that workplace, and cancels from the numerator and denominator of this conditional probability.

The residence probability (2), workplace probability (3) and conditional residence probability (4) play a key role in our empirical analysis below. Using these relationships, we can decompose the observed shares of residents who work where they live for each county ($\lambda_{nnt|n}^{R}$), the share of people who live in each county (λ_{nt}^{R}) and the share of people who work in each county (λ_{nt}^{L}) into the contributions of residence attractiveness ($\mathscr{R}_{nt}$), workplace attractiveness ($\mathscr{W}_{nt}$) and bilateral commuting frictions (κ_{nnt}).

Finally, another key implication of our extreme value specification for idiosyncratic preferences is that expected utility is equalized across all pairs of residences and workplaces:

$$\bar{U}_t = \delta\left(\sum_{r\in N}\sum_{\ell\in N}\left(\mathscr{R}_{rt}\mathscr{W}_{\ell t}/\kappa_{r\ell t}\right)^{\varepsilon}\right)^{1/\varepsilon}, \quad \delta\equiv\Gamma\left(\frac{\varepsilon-1}{\varepsilon}\right), \tag{5}$$

where $\Gamma(\cdot)$ is the Gamma function.

The intuition for this prediction for expected utility is that bilateral commutes with attractive economic characteristics (high workplace and residence attractiveness, and low commuting frictions) attract additional commuters with lower idiosyncratic amenities, until expected utility (taking into account idiosyncratic amenities) is the same across all bilateral commutes.

This specification of workers' commuting decisions in terms of residence attractiveness, workplace attractiveness and bilateral commuting frictions encompasses an entire class of spatial models that are consistent with the structural commuting gravity equation (1). This class of models includes: the classical urban model with one good and no trade costs (as in Ahlfeldt *et al.* 2015)); extensions of the classical urban model with traded and non-traded goods (as in Heblich *et al.* 2020); economic geography versions of the Eaton and Kortum (2002) model with multiple goods and trade costs (as in Redding 2016); economic geography versions of the Armington model with goods differentiated by origin and trade costs (as in Allen and Arkolakis 2014); and new economic geography models with love of variety, increasing returns to scale and trade costs (as in Helpman 1998; Redding and Sturm 2008; Monte *et al.* 2018). Each of the models in this class takes a different stand on what determines residence attractiveness, workplace attractiveness and bilateral commuting frictions. As our quantitative approach below uses only the structural gravity equation (1), it holds throughout this class of models.

II. Data

Our data source is the population census of the US Census Bureau for 1970, 1980, 1990 and 2000. The definition of residence in the population census dates back to the 1790 Census Act and corresponds to 'usual residence', defined as the place where a person lives and sleeps most of the time. The 1960 census was the first to ask place-of-work questions, including the name of the city or town where the work takes place, whether it is inside or outside the city limits, the name of the county, and the name of the state. Beginning with the 1970 census, the place-of-work information was expanded to include the street address and ZIP code of the work location. The definition of workplace corresponds to the place of work in the previous week.

We use publicly available tabulations from the US Census Bureau of the number of bilateral commuters from each workplace county to each residence county for every census decade from 1970 to 2000.[3] We focus on the 48 contiguous US states plus Washington, DC, excluding counties in the states of Alaska and Hawaii. We combine these data with information of the geographical characteristics of counties, including geographical land area and the latitude and longitude of their centroids. We use this information on latitude and longitude to compute the geographical (great circle) distance between the centroids of counties. We use consistent definitions of county boundaries from 1970 to 2000, using 1990 counties as our base and the cross-walk developed in Eckert and Peters (2018). We restrict attention to a balanced panel of counties for which we can construct consistent definitions of county boundaries across all four years. We drop any residence–workplace county pair for which bilateral commuting flows are zero in all four census years from 1970 to 2000.

A key advantage of using counties as our spatial units is that they permit an analysis of changes in commuting patterns over long time horizons. In contrast, commuting flows for more finely detailed spatial units, such as census tracts, are available only for recent years. Although counties are relatively large spatial units compared to census tracts, the commuting network for counties remains relatively dense. In 1970 (2000), the median residence county has 4 (8) workplace counties to which it sends more than 100 commuters, while the median workplace county has 5 (9) residence countries from which it receives more than 100 commuters.

III. Reduced-form Evidence

In this section, we provide reduced-form evidence on the large-scale suburbanization that occurred during the last three decades of the 20th century. We begin by providing evidence on the evolution of the openness of counties to commuting over time. In Figure 1, we display kernel density estimates of the distribution of the share of residents who work in the county where they live ($\lambda^R_{nnt|n} = \lambda_{nnt}/\lambda^R_{nt}$) in 1970 and 2000. At the beginning of our sample period, there is a large concentration of counties with own commuting shares of above 80%, with the median own commuting share equal to 87%. By the end of our sample period, there is marked shift in the distribution of counties towards lower own commuting shares of less than 80%, with the median own commuting share equal to 71%. This pattern is even more marked if we compare the 2000 distribution to the data available on the share of residents who work where they live in the 1960 census, in which the median own commuting share is 91%.[4]

This finding of increased openness to commuting is robust across a wide range of specifications. First, we find a similar pattern using the alternative measure of the own commuting share of the share of workers who live where they work ($\lambda_{nit}/\lambda^L_{it}$), in which the median own commuting share falls from 90% in 1970 to 80% in 2000. Second, in our baseline specification in Figure 1, we focus on unweighted distributions of own commuting shares across counties, as implied by our model and because we are concerned with the distribution of economic activities across space. Nevertheless, we find a similar pattern of increasing openness to commuting if we instead estimate kernel densities, weighting counties by their residence probabilities (λ^R_{nt}) or workplace probabilities (λ^L_{it}). Therefore the typical person at the end of our sample period lives in a county that is more open to commuting than at the beginning of our sample period, with

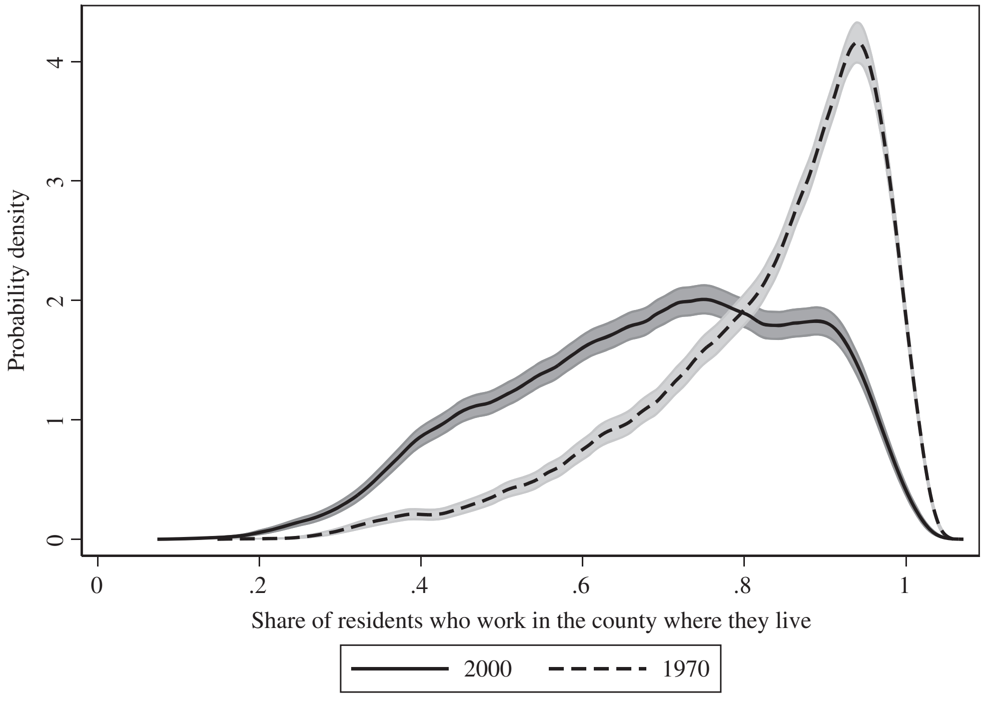

FIGURE 1. Distribution across counties of the share of residents who work in the county where they live in 1970 and 2000 (conditional residence probability, $\lambda^{R}_{nnt|n}$).

Notes: Dashed and solid lines show kernel density estimates of the distribution of the share of residents who work in the county where they live ($\lambda^{R}_{nnt|n} = \lambda_{nnt}/\lambda^{R}_{nt}$) across counties in 1970 and 2000, respectively. Density estimates use the alternative Epanechnikov kernel with optimal (Silverman) bandwidth. Grey shading shows 95% point confidence intervals for $f(x)$, which are constructed as $\widehat{f}_K(x;h) \pm z_{1-\alpha/2}\sqrt{\widehat{V}\{\widehat{f}_K(x;h)\}}$, where h is the bandwidth, $\widehat{V}\{\widehat{f}_K(x;h)\}$ is the exact variance estimate, and $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution. Data from the US population census.

the residence-weighted median county experiencing a fall in the own commuting share from 88% to 76%.

We next turn to the evolution of the workplace employment distribution across counties over time. In Figure 2, we display kernel density estimates of the distribution of the log workplace probability (λ^{L}_{it}) in 1970 and 2000. We find a clear shift in distribution of workplace probabilities towards lower values over time, with a decline in the mass of the distribution at intermediate values, and an increase in the mass in the lower tail. While the point estimates for the upper tail of the distribution of workplace probabilities in 2000 also typically lie above those for 1970, the difference is much smaller and within the 95% confidence intervals. In our baseline specification, we focus on the workplace probability (λ^{L}_{it}), as implied by our model. But we find a similar pattern using the distribution of workplace employment density per unit of geographical land area, with a strong shift in workplace employment towards lower densities over time.[5]

Finally, we examine the evolution of the residence employment distribution across counties over time. In Figure 3, we display kernel density estimates of the distribution of the log residence probability (λ^{R}_{nt}) in 1970 and 2000. Again, we find a shift in the distribution towards lower values over time, with a decline in the mass of the distribution

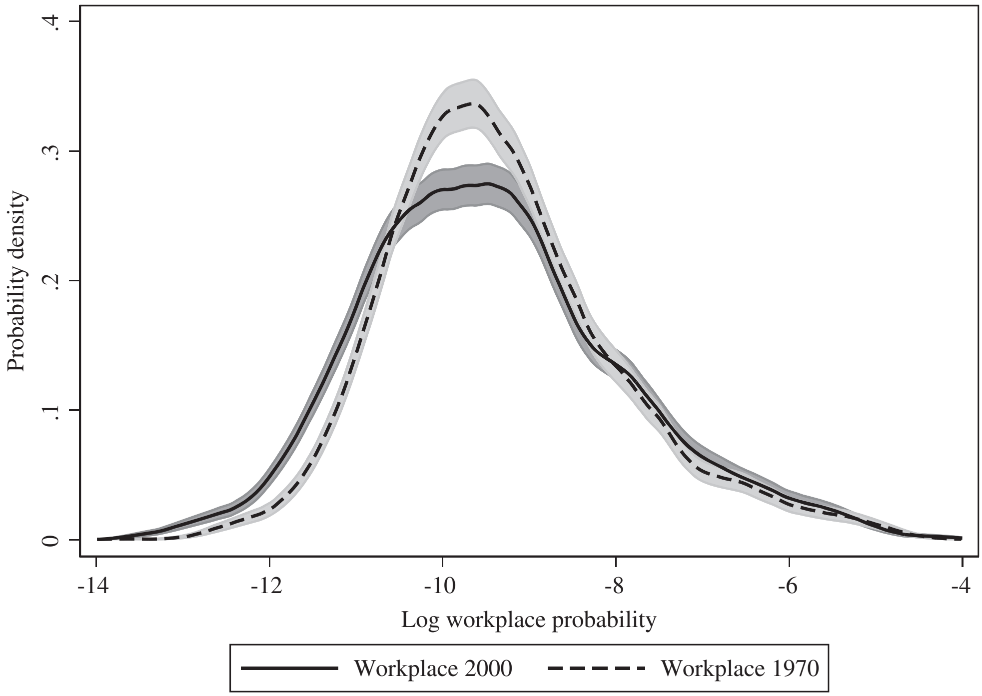

FIGURE 2. Distribution across counties of the share of people who work in each county (workplace probability, λ^L_{nt}).
Notes: Dashed and solid lines show kernel density estimates of the distribution of the share of people who work in each county (λ^L_{nt}) across counties in 1970 and 2000, respectively. Density estimates use the alternative Epanechnikov kernel with optimal (Silverman) bandwidth. Grey shading shows 95% point confidence intervals, which are constructed as in Figure 1. Data from the US population census.

at intermediate values, and an increase in the mass in the lower tail. Although these changes in the distribution of residence employment probabilities are statistically significant at conventional levels, they are smaller in magnitude than those for the distribution of workplace employment probabilities.[6] Therefore while the workplace employment distribution is more spatially concentrated than the residence employment distribution in both 1970 and 2000, it displays a greater shift towards decentralization over time. Again, in our baseline specification in Figure 3, we focus on the residence probability (λ^R_{nt}), as implied by our model. But we find a similar pattern using the distribution of residence employment density per unit of geographical land area, with a shift in residence employment towards lower densities over time.[7]

IV. QUANTITATIVE RESULTS

In this section, we estimate our structural gravity equation for bilateral commuting flows in equation (1), and recover estimates of residence attractiveness ($\mathscr{R}_{nt}$), workplace attractiveness ($\mathscr{W}_{it}$), and bilateral commuting frictions (κ_{nit}). In the first subsection, we provide evidence on the predictive power of the model's structural gravity equation for the observed data. In the second subsection, we construct our estimates of workplace and residence attractiveness for each year. In the final subsection, we undertake our counterfactuals to evaluate the relative contributions of residence attractiveness,

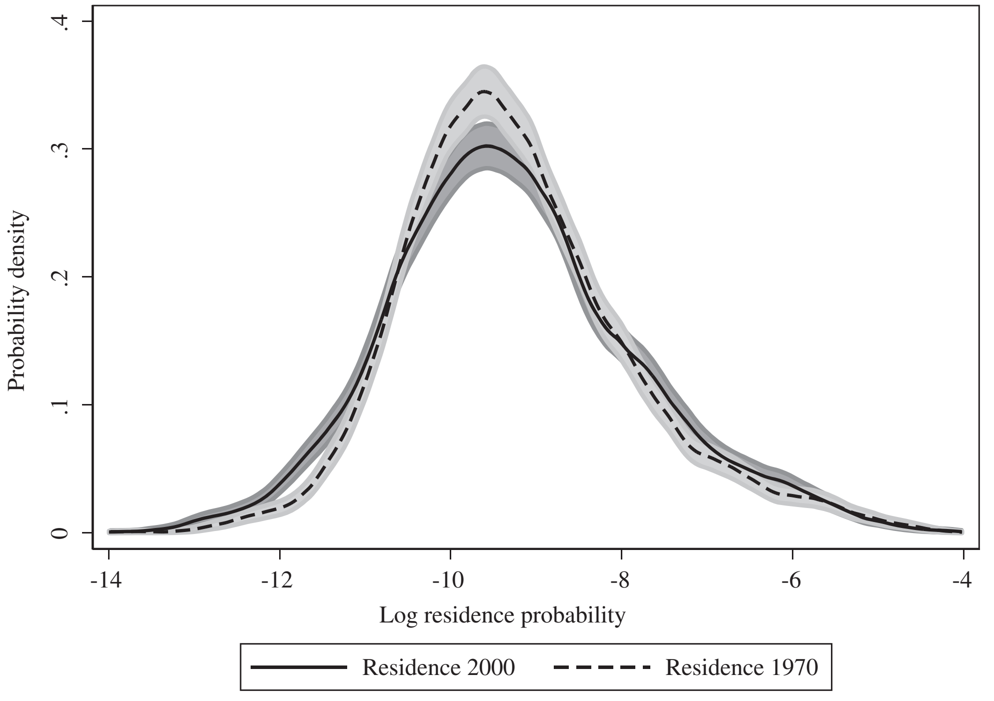

FIGURE 3. Distribution across counties of the share of people who live in each county (residence probability, λ^R_{nt}).
Notes: Dashed and solid lines show kernel density estimates of the distribution of the share of people who live in each county (λ^R_{nt}) across counties in 1970 and 2000, respectively. Density estimates use the alternative Epanechnikov kernel with optimal (Silverman) bandwidth. Grey shading shows 95% point confidence intervals, which are constructed as in Figure 1. Data from the US population census.

workplace attractiveness and bilateral commuting frictions to the observed changes in commuting patterns during our sample period.

Gravity equation estimation

The key insight underlying our approach is to use the observed bilateral commuting data and the log additive structure of the gravity equation in this class of spatial models to reveal the relative importance of these three sets of determinants of the spatial distribution of economic activity. From the unconditional commuting probability (1) and expected utility (5), the probability that a worker commutes from residence n to workplace i at time t can be written as the gravity equation

$$\lambda_{nit} = \chi_t \mathscr{R}^{\varepsilon}_{nt} \mathscr{W}^{\varepsilon}_{it} \operatorname{dist}^{-\delta}_{ni} u_{nit}, \tag{6}$$

where $\mathscr{R}^{\varepsilon}_{nt}$ is a residence fixed effect and $\mathscr{W}^{\varepsilon}_{it}$ is a workplace fixed effect. We have parametrized bilateral commuting frictions between counties as a constant elasticity function of geographical (great circle) distance between their centroids (dist_{ni}) and a stochastic error (u_{nit}): $\kappa^{-\varepsilon}_{nit} = \operatorname{dist}^{-\delta}_{ni} u_{nit}$. The elasticity on distance (δ) is a composite of the elasticity of commuting flows with respect to commuting costs (ε) and the elasticity of commuting costs with respect to distance (φ). The stochastic error u_{nit} captures all components of bilateral commuting frictions that are not captured by bilateral distance, such as transport infrastructure and public transit options. The constant $\chi_t \equiv (\bar{U}_t/\delta)^{-\varepsilon}$ captures the common level of expected utility across all locations, where this constant is separately identified from the fixed effects only up to a normalization.

We estimate the commuting gravity equation (6) for each year separately using the Poisson pseudo maximum likelihood (PPML) estimator of Santos Silva and Tenreyro (2006). This estimator yields theoretically consistent estimates of the fixed effects (as shown in Fally 2015), and allows for granularity and zeros in bilateral commuting flows (as discussed in Dingel and Tintelnot 2020). Separate identification of the residence ($\mathscr{R}_{nt}^{\varepsilon}$) and workplace ($\mathscr{W}_{it}^{\varepsilon}$) fixed effects requires—in the language of graph theory—that the counties included in the estimation sample are connected through commuting networks, either directly or indirectly, which is satisfied in almost all cases.[8]

In Table 1, we report the estimation results, where each column corresponds to a different census year from 1970 to 2000. Consistent with the large reduced-form literature that has estimated commuting gravity equations, we find a negative and statistically significant relationship between bilateral commuting flows and bilateral distance. This estimated coefficient captures both the direct effect of distance on commuting flows and any indirect effect through the provision of less transport infrastructure for commutes of longer distances. Comparing columns (1)–(4), we find that the estimated coefficient on bilateral distance falls over time. This pattern of results is consistent with the idea that the expansion of the interstate highway system and the fall in the real cost of car ownership over our sample period reduced commuting costs over longer distances relative to those over shorter distances.

To explore further the changes in commuting patterns over our sample period, we distinguish between the extensive margin (the number of residence–workplace pairs with positive commuting flows) and the intensive margin (the number of commuters conditional on positive commuting flows). Starting with the extensive margin, of the 57,491 residence–workplace pairs in our sample, we find that 88.3% have positive bilateral commuting flows in 2000, and 85.8% have more than ten commuters in that

TABLE 1
GRAVITY IN COMMUTING 1970-2000

		λ_{nit}		
PPML	(1)	(2)	(3)	(4)
Distance	−3.609***	−3.315***	−3.254***	−3.201***
	(0.110)	(0.081)	(0.075)	(0.066)
Residence fixed effects	Yes	Yes	Yes	Yes
Workplace fixed effects	Yes	Yes	Yes	Yes
Observations	54,296	54,226	54,278	54,393
Pseudo R^2	0.24	0.21	0.20	0.19
Year	1970	1980	1990	2000

Notes
Observations are a cross-section of residence–workplace pairs of counties in a given year. Columns (1)–(4) show results for 1970, 1980, 1990 and 2000, respectively. Distance is the great circle distance (Haversine formula) between the centroids of counties. All specifications include residence fixed effects and workplace fixed effects, and are estimated using the Poisson pseudo maximum likelihood (PPML) estimator. Standard errors in parentheses are heteroscedasticity robust. Data from the US population census.
***denotes significance at the 1% level.

year. As the definition of workplace in the population census is based on the place of work in the previous week, while the definition of residence is based on the place where a person lives and sleeps most of the time, some of these positive observations could reflect business trips. In Figure 4, we display the share of residence–workplace pairs with positive commuting flows in each year, expressed as a share of those with positive commuting flows in 2000, which implies that 2000 mechanically has a share of 1. As is apparent from the figure, we find a more than 40% increase in the number of residence–workplace pairs with positive commuting flows between 1970 and 2000, which occurs gradually over the course of our sample period. Again, this pattern of results is consistent with the progressive expansion of the interstate highway system and the fall in the real cost of car ownership over time reducing commuting costs and raising the number of residence–workplace pairs with positive commuting flows.

Turning now to the intensive margin, we take logs in the commuting gravity equation (6), which drops any zeros from the sample. We next re-estimate this log-linear commuting gravity equation using the two-way fixed effects estimator, and use the Frisch–Waugh–Lovell Theorem to explore the explanatory power of bilateral distance relative to the workplace and residence fixed effects. First, we regress the log unconditional commuting probability on the workplace and residence fixed effects, and generate the residual. Second, we regress the log of bilateral distance on the workplace and residence fixed effects, and generate the residual. Third, we regress the two residuals on one another, which allows us to focus on the conditional correlation between bilateral commuting flows and distance, after conditioning on the workplace and residence fixed effects.

In Figure 5, we display a scatterplot of the two residuals for the year 2000, as well as the linear regression fit between them. We find a strong negative and statistically

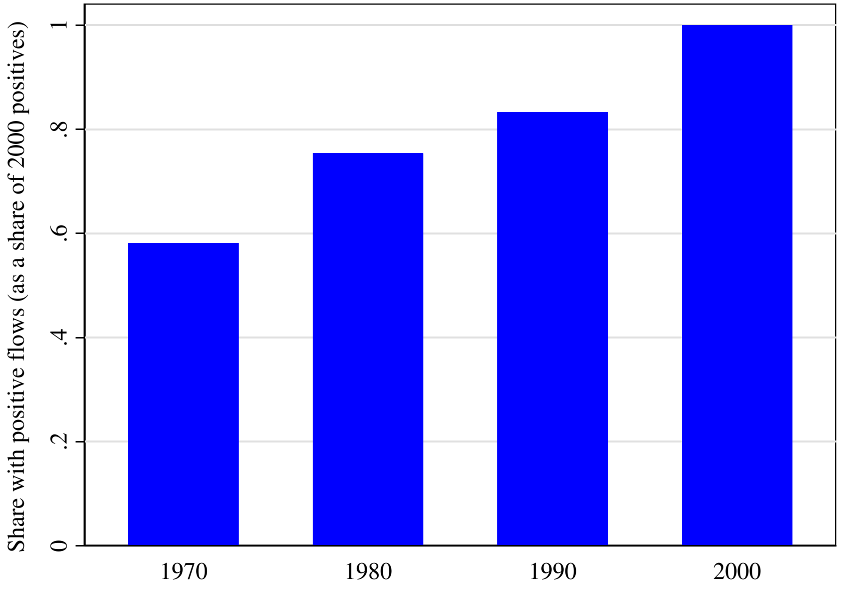

FIGURE 4. Extensive margin of commuting gravity 1970–2000.
Notes: Share of residence–workplace pairs with positive bilateral commuting flows in each year, as a share of the number with positive bilateral commuting flows in the year 2000 (therefore 2000 = 1). Data from the US population census.

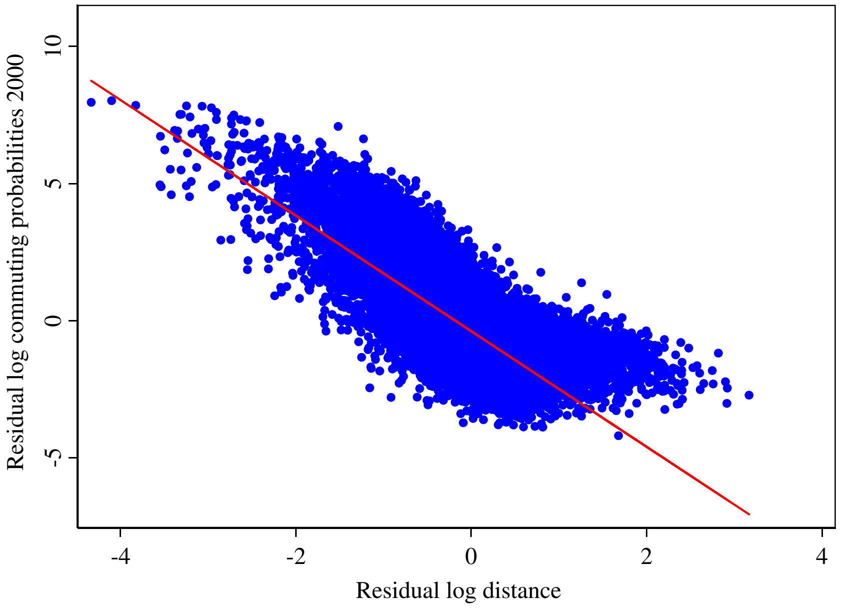

FIGURE 5. Intensive margin of commuting gravity 2000.
Notes: The figure shows residual log commuting probabilities against residual log distance in 2000. Residuals from ordinary least squares regressions of each variable on workplace and residence fixed effects. The straight line shows the linear fit from the regression of the two residuals on one another. Data from the US population census.

significant relationship between residual log commuting flows and residual log distance. We find that residual log distance has substantial explanatory power in this relationship, with an R-squared of 0.55. Consistent with our modelling of bilateral commuting flows as a constant elasticity function of bilateral distance, we find an approximately log-linear relationship between residual log commuting flows and residual log distance. At the very highest values of residual log distance, we observe a slight flattening of this relationship. Recall that the definition of workplace in the population census is based on the place of work in the previous week, while the definition of residence is based on the place where a person lives and sleeps most of the time. Therefore some of these measured commutes over long distances could reflect business trips, which are likely to be less sensitive to bilateral distance. Although, for brevity, we focus on results for the year 2000, we find a similar pattern of results across the other years of our sample.

Estimating workplace and residence attractiveness

We now construct our estimates of workplace and residence attractiveness using the model's predictions for the unconditional commuting probability (1) and expected utility (5). Returning to the gravity equation (6), we drop the term in bilateral distance and absorb all bilateral commuting frictions into the error term:

$$\lambda_{nit} = \mathscr{R}_{nt}^{\varepsilon} \mathscr{W}_{it}^{\varepsilon} \mathscr{K}_{nit}^{-\varepsilon}, \tag{7}$$

where the residence fixed effect ($\mathscr{R}_{nt}^{\varepsilon}$) now captures all characteristics of a residence that affect its commuting flows (including its average distance to other counties), and the workplace fixed effect ($\mathscr{W}_{it}^{\varepsilon}$) now captures all characteristics of a workplace that affect its commuting flows (including its average distance to other counties). The error term ($\mathscr{K}_{nit}^{-\varepsilon} = \chi_t \kappa_{nit}^{-\varepsilon}$) absorbs all bilateral commuting frictions (including bilateral variation in distance), as well as the constant χ_t, where this constant is again separately identified from the fixed effects only up to a normalization.

An advantage of this more parsimonious specification of the commuting gravity equation is that it allows us to estimate residence and workplace fixed effects non-parametrically, without imposing any functional form assumptions on how variables such as bilateral distance affect commuting flows. We can estimate this commuting gravity equation without knowledge of the Fréchet shape parameter ε, which determines the responsiveness of commuting decisions to economic variables in the model, because it is incorporated into our estimates of workplace attractiveness ($\mathscr{W}_{it}^{\varepsilon}$) from the workplace fixed effect, residence attractiveness ($\mathscr{R}_{nt}^{\varepsilon}$) from the residence fixed effect, and bilateral commuting frictions ($\mathscr{K}_{nit}^{-\varepsilon}$) from the error term. In this estimation, we use only our assumptions of structural gravity, namely that bilateral commuting flows depend log additively on residence attractiveness, workplace attractiveness and bilateral commuting frictions. Again, we estimate the gravity equation (7) separately for each year of our sample period. We use the estimated residence fixed effects from equation (7) as our measure of residence attractiveness in each year, the estimated workplace fixed effects from equation (7) as our measure of workplace attractiveness in each year, and the estimated residual from equation (7) as our measure of bilateral commuting frictions.

Counterfactuals

Using our estimates of the structural gravity equation in this class of spatial models, we now undertake counterfactuals to evaluate the relative importance of changes in residence attractiveness, workplace attractiveness and bilateral commuting frictions in explaining the suburbanization observed during our sample period. In the first subsubsection below, we introduce our approach for evaluating the contributions of these three terms to the observed changes in the conditional residence probability ($\lambda_{nnt|n}^{R}$), the residence probability (λ_{nt}^{R}) and the workplace probability (λ_{nt}^{L}). In the second subsubsection, we implement this approach using our bilateral commuting data for US counties from 1970 to 2000.

Exact-hat algebra counterfactuals We undertake our counterfactuals using an 'exact-hat algebra' approach similar to that used in the quantitative international trade literature following Dekle *et al.* (2007). We start at the observed equilibrium in our baseline year of $T = 2000$ at the end of our sample period, and undertake counterfactuals for changes in residence attractiveness ($\mathscr{R}_{nt}^{\varepsilon}$), workplace attractiveness ($\mathscr{W}_{it}^{\varepsilon}$) and bilateral commuting frictions ($\mathscr{K}_{nit}^{-\varepsilon}$), going backwards in time to an earlier year $t < T$. In particular, we use the property of this class of spatial models that the counterfactual equilibrium conditions in any earlier year can be written in terms of the value of the endogenous variables in our baseline year and the relative changes in variables between the two years.

Using this property, we can write the conditional residence probability in an earlier year ($\lambda_{nnt|n}^{R}$) in equation (4) in terms of its value in our baseline year ($\lambda_{nnT|n}^{R}$) and the change in workplace attractiveness ($\widehat{\mathscr{W}}_{nt}^{\varepsilon}$) and bilateral commuting costs ($\widehat{\mathscr{K}}_{nnt}^{-\varepsilon}$) as

$$(8) \qquad \lambda^R_{nnt|n} = \lambda^R_{nnT|n}\hat{\lambda}^R_{nnt|n} = \frac{\lambda^R_{nnT|n}\widehat{\mathscr{W}}^{\varepsilon}_{nt}\widehat{\mathscr{K}}^{-\varepsilon}_{nnt}}{\sum_{\ell \in N}\lambda^R_{n\ell T|n}\widehat{\mathscr{W}}^{\varepsilon}_{\ell t}\widehat{\mathscr{K}}^{-\varepsilon}_{n\ell t}},$$

where we use a hat above a variable to denote a relative change in that variable, such that $\widehat{\mathscr{W}}_{nt} = \mathscr{W}_{nt}/\mathscr{W}_{nT}$.

Similarly, the residence probability in an earlier year (λ^R_{nt}) in equation (2) can be expressed in terms of its value in our baseline year (λ^R_{nT}) and the change in residence attractiveness ($\widehat{\mathscr{R}}^{\varepsilon}_{nt}$) and residents' commuting market access ($\widehat{RMA}^{\varepsilon}_{nt}$) as

$$(9) \qquad \lambda^R_{nt} = \lambda^R_{nT}\hat{\lambda}^R_{nt} = \frac{\lambda^R_{nT}\widehat{\mathscr{R}}^{\varepsilon}_{nt}\widehat{RMA}^{\varepsilon}_{nt}}{\sum_{r \in N}\lambda^R_{rT}\widehat{\mathscr{R}}^{\varepsilon}_{rt}\widehat{RMA}^{\varepsilon}_{rt}},$$

where the change in residents' commuting market access ($\widehat{RMA}^{\varepsilon}_{nt}$) depends on the conditional residence probability in our baseline year ($\lambda^R_{n\ell T|n}$) and the changes in workplace attractiveness ($\widehat{\mathscr{W}}^{\varepsilon}_{it}$) and bilateral commuting costs ($\widehat{\mathscr{K}}_{nit}$):

$$(10) \qquad \widehat{RMA}^{\varepsilon}_{nt} = \sum_{\ell \in N}\lambda^R_{n\ell T|n}\widehat{\mathscr{W}}^{\varepsilon}_{\ell t}\widehat{\mathscr{K}}^{-\varepsilon}_{n\ell t}.$$

Finally, the workplace probability in an earlier year (λ^L_{it}) in equation (3) can be written in terms of its value in our baseline year (λ^L_{iT}) and the changes in workplace attractiveness ($\widehat{\mathscr{W}}^{\varepsilon}_{it}$) and workers' commuting market access ($\widehat{WMA}^{\varepsilon}_{it}$) as

$$(11) \qquad \lambda^L_{it} = \lambda^L_{iT}\hat{\lambda}^L_{it} = \frac{\lambda^L_{iT}\widehat{\mathscr{W}}^{\varepsilon}_{it}\widehat{WMA}^{\varepsilon}_{it}}{\sum_{\ell \in N}\lambda^L_{\ell T}\widehat{\mathscr{W}}^{\varepsilon}_{\ell t}\widehat{WMA}^{\varepsilon}_{\ell t}},$$

where the change in workers' commuting market access ($\widehat{WMA}^{\varepsilon}_{it}$) depends on the conditional workplace probability in our baseline year (λ^L_{iT}) and the changes in residence attractiveness ($\widehat{\mathscr{R}}^{\varepsilon}_{nt}$) and bilateral commuting costs ($\widehat{\mathscr{K}}^{-\varepsilon}_{nit}$):

$$(12) \qquad \widehat{WMA}^{\varepsilon}_{it} = \sum_{r \in N}\lambda^L_{riT|n}\widehat{\mathscr{R}}^{\varepsilon}_{rt}\widehat{\mathscr{K}}^{-\varepsilon}_{rit}.$$

Using the observed initial commuting probabilities in our baseline year of 2000 ($\lambda^R_{nnT|n}$, λ^R_{nT}, λ^L_{iT}) and our estimates of changes in residence attractiveness, workplace attractiveness and bilateral commuting frictions ($\widehat{\mathscr{R}}^{\varepsilon}_{nt}$, $\widehat{\mathscr{W}}^{\varepsilon}_{it}$, $\widehat{\mathscr{K}}^{-\varepsilon}_{nit}$), we can implement the counterfactuals in equations (8)–(12). A number of points about this procedure are worthy of remark.

First, our estimates of residence attractiveness ($\mathscr{R}^{\varepsilon}_{nt}$), workplace attractiveness ($\mathscr{W}^{\varepsilon}_{it}$) and bilateral commuting frictions ($\mathscr{K}^{-\varepsilon}_{nit}$) in each year from equation (7) exactly rationalize the observed commuting flows in each year. Therefore if we start at the observed initial commuting probabilities in our baseline year of $T = 2000$ ($\lambda^R_{nnT|n}$, λ^R_{nT}, λ^L_{iT})

and undertake a counterfactual simultaneously changing all three components of residence attractiveness, workplace attractiveness and bilateral commuting frictions ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$, $\widehat{\mathscr{W}}_{it}^{\varepsilon}$, $\widehat{\mathscr{K}}_{nit}^{-\varepsilon}$), then we necessarily exactly replicate the observed commuting probabilities in an earlier year $t < T$ ($\lambda_{nnt|n}^{R}$, λ_{nt}^{R}, λ_{it}^{L}). We use this property to examine the relative importance of each of these components separately compared to changing them all simultaneously.

Second, our counterfactuals use the observed commuting probabilities in our baseline year of $T = 2000$ ($\lambda_{nnT|n}^{R}$, λ_{nT}^{R}, λ_{iT}^{L}) to control for the determinants of commuting patterns in that baseline year. One implication of this approach is that if the observed commuting probabilities in our baseline year $T = 2000$ are equal to zero, then the counterfactual commuting probabilities in an earlier year $t < T$ are necessarily also equal to zero. An advantage of undertaking our counterfactuals backwards in time is that there are far fewer residence–workplace pairs with positive flows that subsequently become zeros than those with zeros that subsequently become positive flows, because of the more than 40% increase in the number of residence–workplace pairs with positive commuting flows established above.

Third, we use our counterfactuals to evaluate the relative importance of different mechanisms in the model, by changing residence attractiveness ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$), workplace attractiveness ($\widehat{\mathscr{W}}_{iT}^{\varepsilon}$) and bilateral commuting frictions ($\widehat{\mathscr{K}}_{niT}^{-\varepsilon}$) individually, and comparing the predicted commuting probabilities in an earlier year ($\lambda_{nnt|n}^{R}$, λ_{nt}^{R}, λ_{it}^{L}) to the observed values of these commuting probabilities in the data in that earlier year. While we use this exercise to assess the relative importance of different mechanisms, it does not necessarily have a causal interpretation. For example, the change in bilateral commuting frictions ($\widehat{\mathscr{K}}_{nit}^{-\varepsilon}$) itself could be influenced by the changes in residence attractiveness ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$) and workplace attractiveness ($\widehat{\mathscr{W}}_{it}^{\varepsilon}$) that affect the return to endogenous investments in transport infrastructure. Similarly, the changes in residence attractiveness ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$) and workplace attractiveness ($\widehat{\mathscr{W}}_{it}^{\varepsilon}$) themselves could be affected by investments in transport infrastructure that affect the surrounding concentration of economic activity, and hence agglomeration forces. Regardless of the direction in which causality runs, or whether it runs in both directions simultaneously, our counterfactuals isolate the relative importance of these three mechanisms for the observed changes in commuting patterns in the data, as in the literature on growth and business cycle accounting in macroeconomics.

Finally, the counterfactual residence probability (λ_{nt}^{R}) in equation (9) and the counterfactual workplace probability (λ_{nt}^{L}) in equation (11) depend on the changes in all three components of residence attractiveness ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$), workplace attractiveness ($\widehat{\mathscr{W}}_{it}^{\varepsilon}$) and bilateral commuting frictions ($\widehat{\mathscr{K}}_{nit}^{-\varepsilon}$), both directly and through the terms in residents' and workers' commuting market access terms ($\widehat{RMA}_{nt}^{\varepsilon}$ and $\widehat{WMA}_{it}^{\varepsilon}$, respectively). In contrast, the counterfactual conditional residence probability ($\lambda_{nnt|n}^{R}$) in equation (8) depends only on the changes in workplace attractiveness ($\widehat{\mathscr{W}}_{it}^{\varepsilon}$) and bilateral commuting frictions ($\widehat{\mathscr{K}}_{nit}^{-\varepsilon}$), because the changes in residence attractiveness ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$) cancel from the numerator and denominator of this conditional probability, as discussed in Section I above.

Counterfactual predictions We now use our estimates from the commuting gravity equation (7) to implement our three sets of counterfactuals for changes in residence attractiveness ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$), workplace attractiveness ($\widehat{\mathscr{W}}_{it}^{\varepsilon}$) and bilateral commuting frictions ($\widehat{\mathscr{K}}_{nit}^{-\varepsilon}$). We begin with the conditional residence probability ($\lambda_{nnt|n}^{R}$), before turning to the residence probability (λ_{nt}^{R}) and the workplace probability (λ_{nt}^{L}).

Conditional residence probability ($\lambda_{nnt|n}^{R}$) In Figure 6, we reproduce the observed distributions of the share of residents who work in the county where they live in 1970 (medium-dashed line) and 2000 (solid line) from Figure 1 above. Alongside these observed distributions, we show the counterfactual distribution from starting with the observed data in 2000 and changing only bilateral commuting frictions (short-dashed line). We also display the corresponding counterfactual distribution from starting with the observed data in 2000 and changing only workplace attractiveness (long-dashed line). As discussed above, changes in residence attractiveness have no impact on these distributions, because conditional on living in a residence, they do not affect the relative attractiveness of different workplaces.

As is apparent from Figure 6, we find that almost all of the observed increase in the openness of counties to commuting over time is explained by changes in bilateral

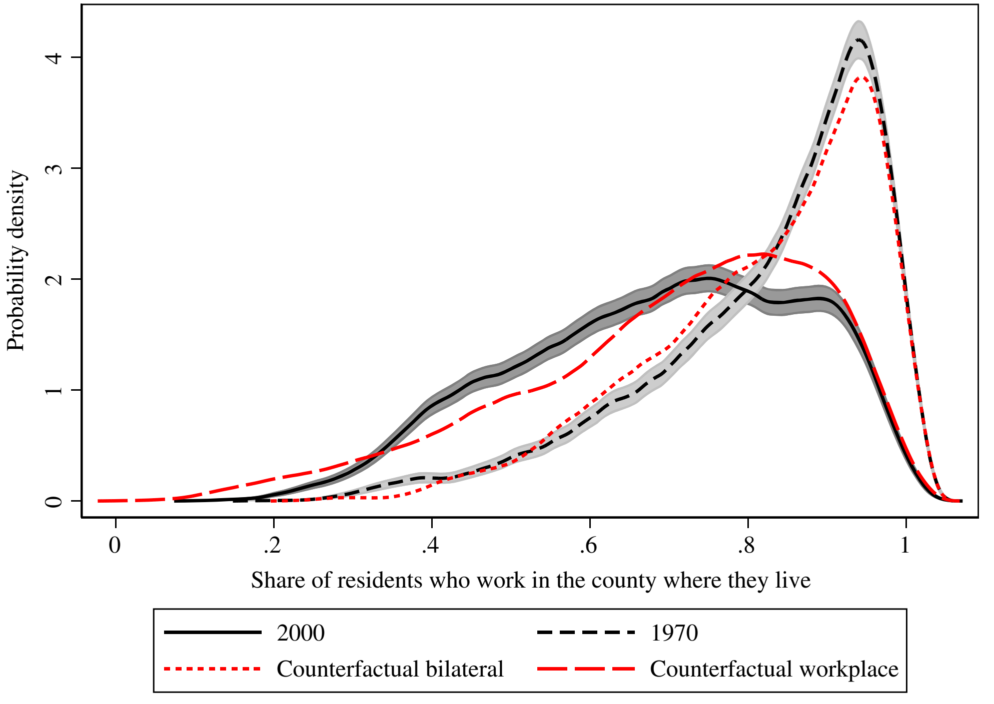

FIGURE 6. Shares of residents who work in the county where they live (conditional residence probability, $\lambda_{nnt|n}^{R}$), actual shares in 2000, and actual and counterfactual shares in 1970.
Notes: Lines show kernel density estimates of the distribution of the share of residents who work in the county where they live ($\lambda_{nnt|n}^{R} = \lambda_{nnt}/\lambda_{nt}^{R}$). The solid line shows values for 2000. The medium-dashed line shows results for 1970. The long-dashed line shows counterfactual results starting at the observed values for 2000 and changing the workplace fixed effects to 1970 values. The short-dashed line shows counterfactual results starting at the observed values for 2000 and changing bilateral commuting frictions to 1970 values. Density estimates use the alternative Epanechnikov kernel with optimal (Silverman) bandwidth. Grey shading shows 95% point confidence intervals, which are constructed as in Figure 1. Data from the US population census.

commuting frictions ($\widehat{\mathscr{K}}_{nit}^{-\varepsilon}$). The counterfactual distribution changing only bilateral commuting frictions closely replicates the sharp increase in the mass of counties at high own commuting shares in the observed 1970 distribution relative to the observed 2000 distribution. We find that changes in workplace attractiveness ($\widehat{\mathscr{W}}_{it}^{\varepsilon}$) do make a contribution to the increased openness of counties to commuting over time. The counterfactual distribution changing only workplace attractiveness has less mass at intermediate values for the own commuting share from 0.4 to 0.6, and greater mass at higher values for the own commuting share from 0.8 to 0.9. Nevertheless, the contribution from changing workplace attractiveness is substantially smaller than that from changing bilateral commuting frictions.

Taken together, this pattern of results is consistent with the view that the primary reason for the increased openness of counties to commuting is reductions in the bilateral costs of commuting, including the expansion of the interstate highway system and the decrease in the real cost of car ownership over time.

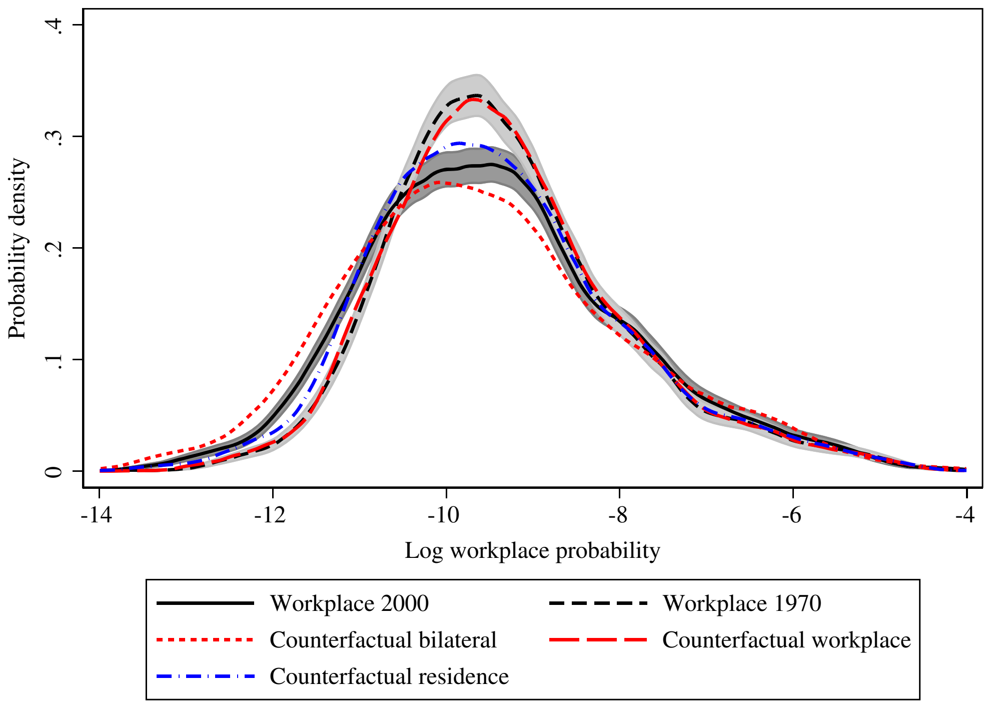

FIGURE 7. Distribution across counties of the share of people who work in each county, actual shares in 2000, actual and counterfactual shares in 1970.

Notes: Lines show kernel density estimates of the distribution of the share of people who work in each county (λ_{nt}^{L}). The solid line shows values for 2000. The medium-dashed line shows results for 1970. The long-dashed line shows counterfactual results starting at the observed values for 2000 and changing the workplace fixed effects to 1970 values. The short-dashed line shows counterfactual results starting at the observed values for 2000 and changing bilateral commuting frictions to 1970 values. The dashed-dotted line shows counterfactual results starting at the observed values for 2000 and changing the residence fixed effects to 1970 values. Density estimates use the alternative Epanechnikov kernel with optimal (Silverman) bandwidth. Grey shading shows 95% point confidence intervals, which are constructed as in Figure 1. Data from the US population census.

Workplace probability (λ_{nt}^{L}) In Figure 7, we reproduce the observed distributions of the share of people who work in each county in 1970 (medium-dashed line) and 2000 (solid line) from Figure 2 above. Alongside these observed distributions, we display three counterfactual distributions. First, we show the counterfactual distribution starting from the observed data in 2000 and changing only bilateral commuting frictions (short-dashed line). Second, we show the counterfactual distribution starting from the observed data in 2000 and changing only workplace attractiveness (long-dashed line). Third, we show the counterfactual distribution starting from the observed data in 2000 and changing only residence attractiveness (dashed-dotted line).

We find that the dominant explanation for the observed changes in the workplace probabilities is workplace attractiveness ($\widehat{\mathscr{W}}_{it}^{\varepsilon}$), which directly affects the choice of workplace in equation (11). The counterfactual distribution changing only workplace attractiveness closely replicates the sharp increase in the mass of counties at intermediate log workplace probabilities (from -10 to -9) and the decrease in the mass of counties at low log workplace probabilities (below -11) in the observed 1970 distribution relative to the observed 2000 distribution.

In contrast, changes in bilateral commuting frictions ($\widehat{\mathscr{K}}_{nit}^{-\varepsilon}$) and residence attractiveness ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$) affect the choice of workplace in equation (11) through the travel-time weighted sum of residents' market access ($\widehat{RMA}_{nt}^{\varepsilon}$). Of these two other determinants of workplace probabilities, we find that changes in residence attractiveness ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$) are somewhat more important than changes in bilateral commuting frictions ($\widehat{\mathscr{K}}_{nit}^{-\varepsilon}$). The counterfactual distribution changing only residence attractiveness generates more of an increase in the mass of counties at intermediate log workplace probabilities (from -10 to -9), as found in moving from the observed 2000 distribution to the observed 1970 distribution. By comparison, the counterfactual distribution changing only bilateral commuting frictions generates too much of an increase in the mass of counties at low log workplace probabilities (below -11), relative to the observed changes in the data.

This pattern of results is consistent with the view that changes in employment opportunities (such as the reallocation of manufacturing away from central cities) are the most important factor in explaining the observed shift in the distribution of workplace probabilities towards lower densities from 1970 to 2000. Changes in residential amenities/disamenities (such as crime and public schools) also play a role. In contrast, as shown in the previous subsubsection, reductions in bilateral commuting probabilities are more important in explaining the increased openness of counties to commuting over time.

Residential probability (λ_{nt}^{R}) In Figure 8, we reproduce the observed distributions of the share of people who live in each county in 1970 (medium-dashed line) and 2000 (solid line) from Figure 3 above. Alongside these observed distributions, we again display three counterfactual distributions, starting from the observed data in 2000 and changing in turn bilateral commuting frictions (short-dashed line), workplace attractiveness (long-dashed line), and residence attractiveness (dashed-dotted line).

We find that changes in residence attractiveness ($\widehat{\mathscr{R}}_{nt}^{\varepsilon}$) and workplace attractiveness ($\widehat{\mathscr{W}}_{it}^{\varepsilon}$) are approximately as important as one another in explaining the observed changes in residence probabilities. Both the counterfactual distribution changing only residence attractiveness and the counterfactual distribution changing only workplace attractiveness show the same increase in the mass of counties at intermediate log residence probabilities (from -10 to -9) and decrease in the mass of counties at low log residence probabilities

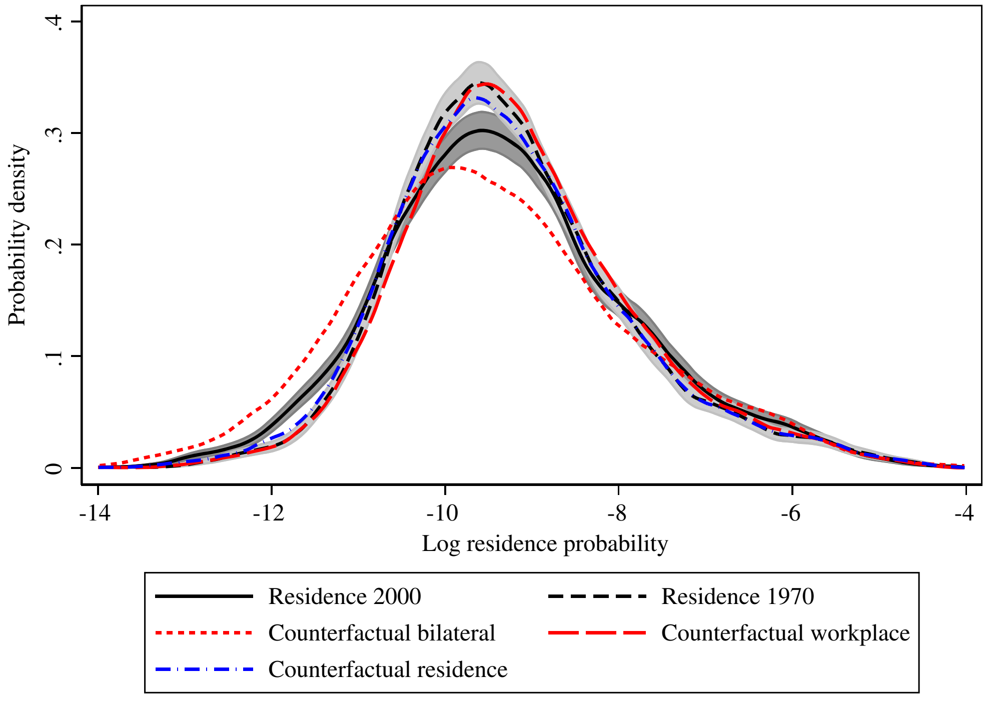

FIGURE 8. Distribution across counties of the share of people who live in each county, actual shares in 2000, actual and counterfactual shares in 1970.
Notes: Lines show kernel density estimates of the distribution of the share of people who live in each county (λ^R_{nt}). The solid line shows values for 2000. The medium-dashed line shows results for 1970. The long-dashed line shows counterfactual results starting at the observed values for 2000 and changing the workplace fixed effects to 1970 values. The short-dashed line shows counterfactual results starting at the observed values for 2000 and changing bilateral commuting frictions to 1970 values. The dashed-dotted line shows counterfactual results starting at the observed values for 2000 and changing the residence fixed effects to 1970 values. Density estimates use the alternative Epanechnikov kernel with optimal (Silverman) bandwidth. Grey shading shows 95% point confidence intervals, which are constructed as in Figure 1. Data from the US population census.

(below −11) as found in moving from the observed 1970 distribution to the observed 2000 distribution. In contrast, the counterfactual distribution changing only bilateral commuting frictions generates too much of an increase in the mass of counties at low log residence probabilities (below −11).

Combined with the results for workplace probabilities above, these findings suggest that changes in employment opportunities (such as the reallocation of manufacturing away from central cities) and residential amenities/disamenities (such as crime and public schools) are important in explaining the overall shift of economic activity towards lower densities from 1970 to 2000. Nevertheless, reductions in bilateral commuting frictions are key to understanding the differential movements in employment by workplace versus employment by residence, and hence the increase in openness to commuting over time.

V. CONCLUSIONS

Suburbanization is one of the most striking features of the USA in the period since the Second World War. Although this decentralization of population and employment is a widely accepted feature of the economy, there remains considerable debate about the economic forces underlying it. One line of research emphasizes workplace forces (such as a dispersion of manufacturing from central cities), while a second group of studies points towards residence factors (amenities/disamenities such as crime and public schools), and a third body of work stresses bilateral commuting costs (including the expansion of the interstate highway system and the falling real cost of car ownership).

We develop a new methodology for discriminating between these three leading explanations for the observed changes in suburbanization. Our methodology holds in an entire class of spatial models that are characterized by a structural gravity equation for commuting, in which bilateral commuting flows depend on bilateral commuting frictions, a workplace fixed effect and a residence fixed effect. The key idea underlying our approach is to use the observed changes in commuting flows and the structure of the gravity equation from this class of models to reveal the relative importance of these different explanations.

We implement this methodology using population census data on bilateral commuting flows between US counties from 1970 to 2000. We document three large-scale changes in observed patterns of commuting in the late 20th century USA. First, counties became substantially more open to commuting flows. Between 1970 and 2000, the median share of residents who work in the county where they live fell from 87% to 71%, and the fraction of counties with values for this share of less than 50% increased almost fourfold from around 5% to about 18%. Second, there was a dispersion of both employment and population from central cities, with the distributions of both workplace and residence employment shifting towards lower densities over this time period. Third, this shift in distribution was larger for workforce employment than for residence employment, with the result that the distribution of workplace employment become substantially less spatially concentrated over time.

We show that changes in bilateral commuting frictions are the dominant force explaining the observed increase in county openness to commuting over time. Holding workplace and residence attractiveness constant at their 2000 values, and changing only bilateral commuting frictions to their 1970 values, the median share of residents who work in the county where they live rises from 71% to 85%—almost as large as the rise to 87% observed in the data. In contrast, we find that changes in residence attractiveness and workplace attractiveness are more important in explaining the shift of employment by workplace and employment by residence towards lower densities over time. This pattern of results is consistent with a role for factors such as the relocation of manufacturing away from central cities and changes in urban amenities in explaining the shift in economic activity towards lower densities. However, reductions in bilateral commuting frictions—for example, through the expansion of the interstate highway system and the falling real cost of car ownership—are central to capturing the observed increase in openness to commuting over time.

ACKNOWLEDGMENTS

I am grateful to Princeton University for research support. This paper was invited for the 100th-year anniversary special issue of *Economica*, published by the London School of Economics. I would like to thank the editors for the invitation to participate. I also would like to thank Henry Overman, Daniel Sturm, an anonymous referee, and participants at the special issue conference for their helpful comments. The US Census Bureau data on commuting from the population census were made available through the NBER Research Project on the Economics of Transportation in the 21st Century, funded by a grant through the National Science Foundation from the US Department of Transportation (see https://www.nber.org/programs-projects/projects-and-centers/transportation-economics-21st-century, accessed 31 July 2021). I would like to thank Nate Baum-Snow and Caitlin Gorback for their help with the data. I am also grateful to Maximilian Schwarz for excellent research assistance. The usual disclaimer applies.

NOTES

1. Source: US Census Bureau, Travel Time to Work in the United States, American Community Survey (ACS) Reports, 2019.
2. In 1960, the median share of residents who work in the county where they live stood even higher, at 91%.
3. These data on bilateral commuting flows were obtained through the NBER Research Project on the Economics of Transportation in the 21st Century, funded by a grant through the National Science Foundation from the US Department of Transportation; see https://www.nber.org/programs-projects/projects-and-centers/transportation-economics-21st-century (accessed 31 July 2021). Although we have data on the share of residents who work in the county where they live from the 1960 population census, we do not have the full matrix of bilateral commuting flows between counties for that year.
4. As our sample includes counties throughout the USA, the observed changes in Figure 1 reflect variation both within and between metro areas, and between metro areas and rural counties. Focusing on counties in metro areas that have five or more counties, and regressing the changes in own commuting shares from 1970 to 2000 on metro area fixed effects, we find an R-squared of 0.12, suggesting that most of the observed changes are driven by variation within metro areas.
5. Again, we find that many of the observed changes in Figure 2 occur within metro areas. Focusing on counties in metro areas that have five or more counties, and regressing log changes in workplace probabilities on metro area fixed effects, we find an R-squared of 0.33.
6. In a Kolmogorov–Smirnov test, we reject the null hypothesis that the 1970 and 2000 distributions of residence probabilities are the same at conventional levels of significance.
7. Again, we find that many of the observed changes in Figure 3 occur within metro areas. Focusing on counties in metro areas that have five or more counties, and regressing log changes in residence probabilities on metro area fixed effects, we find an R-squared of 0.34.
8. We use the ppmlhdfe command for Poisson estimation with high-dimensional fixed effects from Correia *et al*. (2020). Of the more than 54,000 bilateral workplace–residence pairs in the regression sample, fewer than 200 singleton observations are dropped in each year.

REFERENCES

AHLFELDT, G., REDDING, S., STURM, D. and WOLF, N. (2015). The economics of density: evidence from the Berlin Wall. *Econometrica*, **83**, 2127–89.

ALLEN, T. and ARKOLAKIS, C. (2014). Trade and the topography of the spatial economy. *Quarterly Journal of Economics*, **129**, 1085–140.

ALLEN, T., ARKOLAKIS, C. and LI, X. (2017). Optimal city structure. Mimeo, Yale University.

ALONSO, W. (1964). *Location and Land Use*. Cambridge, MA: Harvard University Press.

ANGEL, S. and LAMSON-HALL, P. (2014). The rise and fall of Manhattan's densities 1800–2010. Marron Institute of Urban Management Working Paper no. 18.

BAUM-SNOW, N. (2007). Did highways cause suburbanization? *Quarterly Journal of Economics*, **122**, 775–80.

BAUM-SNOW, N. (2020). Urban transport expansions and changes in the spatial structure of US cities: implications for productivity and welfare. *Review of Economics and Statistics*, **102**(5), 929–45.

BAUM-SNOW, N., BRANDT, L., HENDERSON, J., TURNER, M. A. and ZHANG, Q. (2017). Roads, railroads and decentralization of Chinese cities. *Review of Economics and Statistics*, **99**, 435–48.

BAUM-SNOW, N., HENDERSON, V., TURNER, M. A., ZHANG, Q. and BRANDT, L. (2020). Does investment in national highways help or hurt hinterland city growth? *Journal of Urban Economics*, **115**, 103–24.

BAUM-SNOW, N. and KAHN, M. E. (2005). Effects of urban rail transit expansions: evidence from sixteen cities, 1970–2000. *Brookings–Wharton Papers on Urban Affairs*, 147–97.

BENABOU, R. (1996). Heterogeneity, stratification, and growth: macroeconomic implications of community structure and school finance. *American Economic Review*, **86**, 584–609.

BENABOU, R. and FERNANDEZ, R. (1996). Income distribution, communities, and the quality of public education. *Quarterly Journal of Economics*, **111**, 135–64.

BILLINGS, S. B. (2011). Estimating the value of a new transit option. *Regional Science and Urban Economics*, **41**, 525–36.

BOUSTAN, L. P. (2000). Was postwar suburbanization 'white flight'? Evidence from the black migration. *Quarterly Journal of Economics*, **125**, 417–43.

BOUSTAN, L. P. and MARGO, R. A. (2009). Job decentralization and residential location. *Brookings–Wharton Papers on Urban Affairs*, 1–31.

BREZIS, E. S. and KRUGMAN, P. R. (1997). Technology and the life cycle of cities. *Journal of Economic Growth*, **2**, 369–83.

BROOKS, L. and LUTZ, B. (2018). Vestiges of transit: urban persistence at a micro scale. *Review of Economics and Statistics*, **101**, 385–99.

CALIENDO, L., PARRO, F., ROSSI-HANSBERG, E. and SARTE, P.-D. (2018). The impact of regional and sectoral productivity changes on the U.S. economy. *Review of Economic Studies*, **85**, 2042–96.

CHANDRA, A. and THOMPSON, E. (2000). Does public infrastructure affect economic activity? Evidence from the rural interstate highway system. *Regional Science and Urban Economics*, **30**, 457–90.

CHETTY, R., HENDREN, N., JONES, M. R. and PORTER, S. R. (2020). Race and economic opportunity in the United States: an intergenerational perspective. *Quarterly Journal of Economics*, **135**, 711–83.

COMBES, P.-P. and GOBILLON, L. (2015). The empirics of agglomeration economies. In G. Duranton, J. V. Henderson and W. C. Strange (eds), *Handbook of Regional and Urban Economics*, Vol. **5A**. Amsterdam: Elsevier North-Holland, pp. 247–348.

CORREIA, S., GUIMARAES, P. and ZYLKIN, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. *Stata Journal*, **20**, 95–115.

COUTURE, V., GAUBERT, C., HANDBURY, J. and HURST, E. (2020). Income growth and the distributional effects of urban spatial sorting. University of California, Berkeley.

COUTURE, V. and HANDBURY, J. (2020). Urban revival in America. *Journal of Urban Economics*, **119**.

DAVIS, D. R. and WEINSTEIN, D. E. (2002). Bones, bombs, and break points: the geography of economic activity. *American Economic Review*, **92**, 1269–89.

DEKLE, R., EATON, J. and KORTUM, S. (2007). Unbalanced trade. *American Economic Review*, **97**, 351–5.

DESMET, K. and ROSSI-HANSBERG, E. (2009). Spatial growth and industry age. *Journal of Economic Theory*, **144**, 2477–502.

DINGEL, J. I. and TINTELNOT, F. (2020). Spatial economics for granular settings. NBER Working Paper no. 27287.

DONALDSON, D. (2018). Railroads of the Raj: estimating the impact of transportation infrastructure. *American Economic Review*, **108**, 899–934.

DONALDSON, D. and HORNBECK, R. (2016). Railroads and American economic growth: a market access approach. *Quarterly Journal of Economics*, **131**, 799–858.

DURANTON, G., MORROW, P. and TURNER, M. A. (2014). Roads and trade: evidence from the US. *Review of Economic Studies*, **81**, 681–724.

DURANTON, G. and PUGA, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson and J.-F. Thisse (eds), *Handbook of Regional and Urban Economics*, Vol. **4**. Amsterdam: Elsevier, pp. 2063–117.

DURANTON, G. and TURNER, M. A. (2011). The fundamental law of road congestion: evidence from US cities. *American Economic Review*, **101**, 2616–52.

DURANTON, G. and TURNER, M. A. (2012). Urban growth and transportation. *Review of Economic Studies*, **79**, 1407–40.

EATON, J. and KORTUM, S. (2002). Technology, geography, and trade. *Econometrica*, **70**, 1741–79.

ECKERT, F. and KLEINEBERG, T. (2021). Saving the American dream? Education policies in spatial general equilibrium. Mimeo, University of California, San Diego.

ECKERT, F. and PETERS, M. (2018). A crosswalk for US spatial data since 1790. Mimeo, Yale University.

FABER, B. (2014). Integration and the periphery: the unintended effects of new highways in a developing country. *Review of Economic Studies*, **81**, 1046–70.

FALLY, T. (2015). Structural gravity and fixed effects. *Journal of International Economics*, **97**, 76–85.

FOGELSON, R. M. (2003). *Downtown: Its Rise and Fall, 1880–1950*. New Haven, CT: Yale University Press.

FOGLI, A. and GUERRIERI, V. (2019). The end of the American dream? Inequality and segregation in US cities. NBER Working Paper no. 26143.

FOTHERINGHAM, A. S. and O'KELLY, M. E. (1989). *Spatial Interaction Models: Formulations and Applications*. Dordrecht: Kluwer.

FUJITA, M. and KRUGMAN, P. (1995). When is the economy monocentric? Von Thünen and Chamberlain unified. *Regional Science and Urban Economics*, **25**, 505–28.

FUJITA, M., KRUGMAN, P. and VENABLES, A. J. (1999). *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.

FUJITA, M. and OGAWA, H. (1982). Multiple equilibria and structural transformation of nonmonocentric urban configurations. *Regional Science and Urban Economics*, **12**, 161–96.

FUJITA, M. and THISSE, J.-F. (2002). *Economics of Agglomeration: Cities, Industrial Location and Regional Growth*. Cambridge: Cambridge University Press.

GIBBONS, S. and MACHIN, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics*, **57**, 148–69.

GLAESER, E. L. and KOHLHASE, J. E. (2004). Cities, regions and the decline of transport costs. *Papers in Regional Science*, **83**, 197–228.

GLAESER, E. L. and SACERDOTE, B. (1999). Why is there more crime in cities? *Journal of Political Economy*, **107**, S225–S258.

GUERRIERI, V., HARTLEY, D. and HURST, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, **100**, 45–60.

HEBLICH, S., REDDING, S. J. and STURM, D. M. (2020). The making of the modern metropolis: evidence from London. *Quarterly Journal of Economics*, **135**, 2059–133.

HELPMAN, E. (1998). The size of regions. In D. Pines, E. Sadka and I. Zilcha (eds), *Topics in Public Economics: Theoretical and Applied Analysis*. Cambridge: Cambridge University Press, pp. 33–54.

HENDERSON, J. V. (1974). The sizes and types of cities. *American Economic Review*, **64**, 640–56.

HERSHBERG, T. (1981). *Philadelphia: Work, Space, Family, and Group Experience in the Nineteenth Century*. Oxford: Oxford University Press.

HORNBECK, R. and ROTEMBERG, M. (2021). Railroads, market access, and aggregate productivity growth. Mimeo, Chicago Booth School of Business.

JACKSON, K. T. (1987). *Crabgrass Frontier: The Suburbanization of the United States*. Oxford: Oxford University Press.

JACOBS, J. (1992). *The Death and Life of Great American Cities*. New York: Vintage Books.

KLEINMAN, B., LIU, E. and REDDING, S. (2021). Dynamic spatial general equilibrium. Centre for Economic Performance Discussion Paper no. 1785.

KLINE, P. and MORETTI, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics*, **129**, 275–331.

KRUGMAN, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, **99**, 483–99.

KRUGMAN, P. and VENABLES, A. (1995). Globalization and the inequality of nations. *Quarterly Journal of Economics*, **110**, 857–80.

LEE, S. K. (2020). Crabgrass Frontier revisited in New York: through the lens of 21st-century data. Mimeo, Yale University.

LUCAS, R. E. and ROSSI-HANSBERG, E. (2002). On the internal structure of cities. *Econometrica*, **70**, 1445–76.

MCDONALD, J. F. and MCMILLEN, D. P. (2010). *Urban Economics and Real Estate: Theory and Policy*. Hoboken, NJ: John Wiley & Sons.

MCDONALD, J. F. and OSUJI, C. I. (1995). The effect of anticipated transportation improvement on residential land values. *Regional Science and Urban Economics*, **25**, 261–78.

MCFADDEN, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, **3**, 303–28.

MICHAELS, G. (2008). The effect of trade on the demand for skill: evidence from the interstate highway system. *Review of Economics and Statistics*, **90**, 683–701.

MICHAELS, G., RAUCH, F. and REDDING, S. (2019). Task specialization in U.S. cities from 1880–2000. *Journal of the European Economic Association*, **17**, 754–98.

MILLS, E. S. (1967). An aggregative model of resource allocation in a metropolitan centre. *American Economic Review*, **57**, 197–210.

MIYAUCHI, Y., NAKAJIMA, K. and REDDING, S. (2021). Consumption access and agglomeration: evidence from smartphone data. CEPR Discussion Paper no. 6741-1613580728.

MONTE, F., REDDING, S. and ROSSI-HANSBERG, E. (2018). Commuting, migration and local employment elasticities. *American Economic Review*, **108**, 3855–90.

MORETTI, E. (2011). Local labor markets. In D. Card and O. Ashenfelter (eds), *Handbook of Labor Economics*, Vol. 4b. Amsterdam: Elsevier North-Holland, pp. 1238–303.

MUTH, R. (1969). *Cities and Housing*. Chicago, IL: University of Chicago Press.

OWENS, R. III, ROSSI-HANSBERG, E. and SARTE, P.-D. (2020). Rethinking Detroit. *American Economic Journal*, **12**, 258–305.

RAE, D. W. (2005). *City: Urbanism and Its End*. New Haven, CT: Yale University Press.

REDDING, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics*, **101**, 148–67.

REDDING, S. J. (2021). Trade and geography. In G. Gopinath, E. Helpman and K. Rogoff (eds), *Handbook of International Economics*, Vol. **5**. Amsterdam: Elsevier, forthcoming.

REDDING, S. J. and ROSSI-HANSBERG, E. (2017). Quantitative spatial models. *Annual Review of Economics*, **9**, 21–58.

REDDING, S. J. and STURM, D. M. (2008). The costs of remoteness: evidence from German division and reunification. *American Economic Review*, **98**, 1766–97.

REDDING, S. J. and TURNER, M. A. (2015). Transportation costs and the spatial organization of economic activity. In G. Duration, J. V. Henderson, and W. Strange (eds), *Handbook of Regional and Urban Economics*. Amsterdam: Elsevier, pp. 1339–98.

ROSENTHAL, S. S. and STRANGE, W. C. (2004). Evidence on the nature and sources of agglomeration economics. In J. V. Henderson and J. Thisse (eds), *Handbook of Regional and Urban Economics*, Vol. **4**. Amsterdam: North-Holland.

ROSSI-HANSBERG, E. (2005). A spatial theory of trade. *American Economic Review*, **95**, 1464–91.

SANTOS SILVA, J. M. C. and TENREYRO, S. (2006). The log of gravity. *Review of Economics and Statistics*, **88**, 641–58.

STEVENS, J. J. and HOLMES, T. J. (2004). Spatial distribution of economic activities in North America. J. V. Henderson and J. Thisse (eds), *Handbook of Regional and Urban Economics*, Vol. **4**. Amsterdam: North-Holland, pp. 2797–843.

TSIVANIDIS, N. (2018). The aggregate and distributional effects of urban transit infrastructure: evidence from Bogotá's TransMilenio. Mimeo, University of California, Berkeley.

WARNER, S. B. (1978). *Streetcar suburbs: the process of growth in Boston, 1870–1900*, Cambridge, MA: Harvard University Press.