University of Pennsylvania
ScholarlyCommons

Wharton Research Scholars
Wharton Undergraduate Research

5-2019

Neighborhood Loyalty or Neighborhood Entrapment? Explaining Unmeasured Sources of Reduced Geographic Mobility

Prakash Mishra
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/wharton_research_scholars

Part of the Business Commons, and the Urban Studies and Planning Commons

Mishra, Prakash, "Neighborhood Loyalty or Neighborhood Entrapment? Explaining Unmeasured Sources of Reduced Geographic Mobility" (2019). *Wharton Research Scholars*. 180.
https://repository.upenn.edu/wharton_research_scholars/180

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/wharton_research_scholars/180
For more information, please contact repository@pobox.upenn.edu.

Neighborhood Loyalty or Neighborhood Entrapment? Explaining Unmeasured Sources of Reduced Geographic Mobility

Abstract

This paper develops a simple self-selection utility model for leaving a neighborhood. This opens the door for a simple reduced form approach that leverages a hierarchical Bayesian model to obtain an annualized latent push and pull factor for each neighborhood. Posterior analysis indicates that common predictors of neighborhood quality and inputs to classical utility functions systematically under-predict the number of people who stay in a neighborhood. Such underprediction of out-migration can either be viewed as an unexplained variation due to neighborhood \loyalty" or as financial barriers to mobility. I isolate this residual, referred to as an \inertia," isolated using a quasi-experimental matching method that uses variation in push factor to isolate eects on pull factors. I show the inertia measure can be explained by measures of financial access including distance to a bank branch, local rates of second mortgage, and redevelopment certifications. The residual from these financial measures is shown to correlate with existence of anchor institutions like charter schools. This methodology creates a robust measure of not only the local push- and pull-factors by neighborhood, but also is suggestive of an economic approach to appraising local community strength.

Keywords

neighborhood studies, gentrification, geographic mobility

Disciplines

Business | Urban Studies and Planning

This thesis or dissertation is available at ScholarlyCommons: https://repository.upenn.edu/wharton_research_scholars/180

Neighborhood Loyalty or Neighborhood Entrapment? Explaining Unmeasured Sources of Reduced Geographic Mobility

By

Prakash Mishra

An Undergraduate Thesis submitted in partial fulfillment of the requirements for the

Wharton Research Scholars

Faculty Advisor:

Robert Inman

Richard King Mellon Professor of Finance

Professor of Business Economics & Public Policy

THE WHARTON SCHOOL, UNIVERSITY OF PENNSYLVANIA

MAY 2019

Neighborhood Loyalty or Neighborhood Entrapment? Explaining Unmeasured Sources of Reduced Geographic Mobility

Prakash Mishra
University of Pennsylvania
email: mishrap@wharton.upenn.edu
Advised by Dr. Robert Inman

April 16, 2019

Abstract

This paper develops a simple self-selection utility model for leaving a neighborhood. This opens the door for a simple reduced form approach that leverages a hierarchical Bayesian model to obtain an annualized latent push and pull factor for each neighborhood. Posterior analysis indicates that common predictors of neighborhood quality and inputs to classical utility functions systematically under-predict the number of people who stay in a neighborhood. Such under-prediction of out-migration can either be viewed as an unexplained variation due to neighborhood "loyalty" or as financial barriers to mobility. I isolate this residual, referred to as an "inertia," isolated using a quasi-experimental matching method that uses variation in push factor to isolate effects on pull factors. I show the inertia measure can be explained by measures of financial access including distance to a bank branch, local rates of second mortgage, and redevelopment certifications. The residual from these financial measures is shown to correlate with existence of anchor institutions like charter schools. This methodology creates a robust measure of not only the local push- and pull-factors by neighborhood, but also is suggestive of an economic approach to appraising local community strength.

1 Introduction

Residents move out of neighborhoods when they find new jobs, find better housing or schools, or seek a change in their built environment. Put in other terms, the decision to leave the current neighborhood hinges on the neighborhood no longer serving a demand or need for a particular amenity, given its cost of living. Residents of a neighborhood value the amenities provided by their neighborhood, such as employment access, public goods, or structural access to consumption goods. However, different residents value different amenities with varying importances, and these relative amenity values can further vary over time. To the experimenter looking at economic data, a resident only reveals their preferences for individual amenities when they enter or leave a neighborhood. We cannot observe their taste for individual components of a neighborhood, but when they choose to leave a neighborhood, we know they wanted a different of a given component for the same cost. This is the logic of the neighborhood choice model.

While neighborhood choice has been a key asset to the amenity valuation literature, the process of leaving neighborhoods is perhaps less consistently used for this purpose, though its logic is relatively similar. I demonstrate that a Bayesian survival analysis approach to residential out-migration can be used to estimate demand for amenities. I also demonstrate that such models tend to under-predict the frequency of high-retention neighborhoods - in other words, some residents appear to stay in neighborhoods despite a measurably under-performing basket of amenities given local cost of living. I investigate two possible reasons for what I will refer to as residential inertia.

The first source of inertia is a fixed cost. Leaving a neighborhood involves a large financial shift: accruing the resources to conduct repairs to make a house marketable, acquiring resources to market the home, finding a buyer, finding a new home and gainful employment in a new neighborhood, and the finances for a new down payment. Even renters face a cost of search and movement, as well as potential fees associated with early lease termination or unpaid collections. The complex set of costs creates a cost of geographic mobility that can change the resettlement calculus. This cost of resettlement can be restrictive, especially when trying to leave neighborhoods whose land values have fallen during the time of residency. Such a combination of forces creates a financial quagmire which can trap of individuals in neighborhoods whose amenities are functioning below most measures of market clearing rates. These fixed costs pose problems for econometricians, as they are difficult to measure without targeted data, and often vary with neighborhood due to spatial variation in access to local credit (Tranfaglia, 2018) and faith in credit institutions. (Blanchflower et al., 2003)

The second source of inertia is community. Some residents might choose to stay in an otherwise under-performing neighborhood because there are strong community effects in place. Moving would require high start-up costs for getting such a community in a new location, and this diminishes the potential utility of new neighborhoods.

This paper will investigate geographic mobility by modeling the underlying decision to move out of a neighborhood. While neighborhood choice models tend to be expensive, I use a machine learning approach that learns latent "push" and "pull" factors of neighborhoods, parameters that describe a neighborhood's opportunity cost and amenity value, respectively. By using when residents leave a neighborhood, I construct a distribution of these possible push and pull factors for the population of residents. Based on these factors, I can then construct elasticities of demand with respect to time-varying amenities and time-varying prices. The model is demonstrated to over-predict out-migration from neighborhoods, which I argue is due to a residential inertia factor. In other words, people stay in neighborhoods that do not provide rational benefit for their relative cost.

2 Literature

2.1 Neighborhood Choice

A number of studies have looked at the neighborhood as a basket of goods in the past. Most common is the tendency to study neighborhood choice as a discrete choice experiment; the new resident has a set of different baskets to choose from, and must choose one. Such a framework allows the economist to extract a willingness to pay for different amenities using the basic principles of a dynamic model. (Bayer et al., 2016) The underlying variation in amenities serves as a de facto natural experiment for the economist, with neighborhood selection revealing preference of the consumer for what amenities have a higher value. Others use more traditional hedonic models for a similar purpose, but these models have come under scrutiny for being far too subject to sample choice, especially without motivating structural justifications. (Huang et al., 2014) In the most basic structural models, individuals choose community j to maximize

$$\max_j V(\alpha, g_j, p_j, y)$$

where the inputs are a taste parameter, a congestable public good, housing price, and income in this order. Notice that such inputs are distributional variables; the value function framework naturally translates into a Bayesian construction. A good deal of work on the analogue, the Dirichlet model for multi-modal choice, has been conducted by economists studying conjoint analysis methods for revealed preference. (Louviere, 2008)(Morikawa et al., 2002)

Structural estimation, of course, requires significant overhead in terms of data required and specification of the statistical model Equilibrium existence is difficult to prove deterministically for situations which add particularly complex upgrades to the basic specification of the value function. The nested fixed-point algorithm is most often used to solve for a particular maximization problem. Such a dynamic programming approach can thus be expensive, and partial solution approaches are idiosyncratic and application-specific.

Regardless, significant strides have been made in developing fairly robust forms of these models. Patrick Bayer and Fernando Ferreira in particular have used classic value function iteration methodology to tackle high-risk lendings impact on racial lines in mortgage rates. (Bayer et al., 2017) Bayer also developed a dynamic model earlier on which flexibly calculated marginal willingness to pay for various amenities in an expansive dataset from San Francisco which has since led to a number of augmentations and applications to other regions of the country. (Bayer et al., 2016) Recently, a working paper identifying central neighborhood change in cities developed a modification of the framework to look at a specific piece of the American city over time. (Baum-Snow and Hartley, 2017) Research on housing vouchers has been an important policy area in which this econometric methodology has made recent headway. (Davis et al., 2017)

2.2 Stability in Local Income Distributions

Throughout this work, Lee and Lin (2017) serves as one of several key references because it explicitly tracks the relative stability of a neighborhood's income distribution. Lee & Lin tie part of the variation in income distribution stability to natural amenities. They define an endogenous aggregate amenity level A as:

$$A_{j,t} \equiv \alpha_j + E(\theta|j,t) + m_t + \epsilon_{j,t} \tag{1}$$

α_j is the persistent natural amenity value of neighborhood j (across periods). Then, $E(\theta|j,t)$ is the average income of neighborhood j in period j - meant to capture amenities like school quality which tend to correlate with income. m_t captures trends in city-level amenities shared across neighborhoods. The key to the model in Lin's paper, and the focus of the model in this paper, is $\epsilon_{j,t}$. This is the idiosyncratic amenity shock term, corresponding to natural disasters and governance changes in Lin's paper. The Lee and Lin (2017) paper develops equilibrium conditions and structurally parametrizes neighborhoods which change income-distribution rankings across periods. The paper's key result is that naturally heterogeneous cities have persistent spatial distributions of income, where persistence can be measured as the expected variance of an neighborhood's income percentile rank: $E[\text{Var}(r_{j,t}|j)]$.

A larger discussion exists around neighborhoods which experience a persistent level of income inequality. Benabou (2000) argues in the optimal case such a feature of the economy at its core violates a notion of social justice via the social contract. This discussion largely began with the work of Durlauf in his 1996 paper, which develops persistent income inequality as an artifact of the persistence of income within families. (Durlauf, 1996) This is the obvious candidate; transmission of wealth underpins intergenerational transmission of income characteristics. Barakova et al. (2003) found that credit was indeed an important factor in the tenureship decision. Important also is whether that wealth stays in one place, this being the subject of the current paper.

2.3 Geographic Mobility

Geographic mobility has not been heavily studied amongst economists but bears important relevance to how well current models apply to path-dependent housing markets. Other sources of path dependence have been explored (Bleakley and Lin, 2012), though the mobility path is less explicitly explored. For example, recent work has explored down payments as a key mechanism which stymies homeownership of new homeowners. The out-migration pathway is less explicitly explored. (Stein, 1995)

Famous studies of mobility tend to be on the front of employment or intergenerational socioeconomic mobility. (Chetty et al., 2014) (Chetty et al., 2017) This paper observes the intra-generational mobility problem; fundamentally, we ask:

If people live in a place that limits their economic potential, can they leave it? Compare this to the conventional economic picture of intergenerational income mobility; the two are fundamentally different processes but also fundamentally linked. Evidence exists that economic mobility is tied to spatial features. For example, urban sprawl seems to correlate with lower rates of urban mobility, attributed significantly to ease of job accessibility. (Ewing et al., 2016) Then, leaving a neighborhood from which employment is inaccessible is a fundamental functional necessity in the economic mobility puzzle. Individuals in bottom-quartile neighborhoods are estimated to be able to earn over $600,000 more in lifetime wages if raised in top-quartile neighborhoods, so that neighborhood income has roughly half the effect on future earnings as parental income. (Rothwell and Massey, 2015) With such a large economic impetus to move, the reasons why individuals cannot move, tied to affordability of other neighborhoods or the cost of leaving one's own neighborhood along various dimensions, is a relevant economic question. For example, often, once an individual has made the decision to move, other neighborhoods which are affordable may exist, but outside the city. Costs of re-establishing community, social networks, job connections, childhood schooling relationships, and more can quickly daunt potential movers. (Durlauf, 1993) (Brock and Durlauf, 2001) (Ioannides and Datcher Loury, 2004) (Benabou, 1996)

2.4 Probabilistic Approaches

The analysis in this paper aligns with a growing literature on probabilistic approaches to neighborhood choice analysis. There are several chief reasons for such a methodology. The data most available and most widely available to policymakers consists of aggregates at particular levels, making either a Hierarchical Bayesian aggregate approach or an agent-based approach attractive. Individuals across fields recognize the value of probabilistic approaches for conjoint analysis (Halme and Kallio, 2014) and for understanding nonstationary processes. (Fader and Lattin, 1993)

Recent work has pinpointed gentrification as a fairly nonstationary process overall, described as a timing problem on a neighborhood. Such timing processes difficult for the dynamic programming model, as they require significant overhead to parametrize precisely. (Hwang and Lin, 2016) (Lin et al., 2017) Gentrification and disinvestment are functions of consumer tastes and shifts in market tendencies which may be better captured by hierarchical distributions than deterministic settings. Urban revival mandates across cities have also presented a policy-level need for more sophisticated modeling of how neighborhood taste drifts over time. (Couture and Handbury, 2017)

The Bayesian approach has gained currency in other fields of economics as a form of a discrete choice model which, though perhaps less able to discern between supply-side and demand-side effects in its current form, is perhaps better able to capture distribution-level drift. (Athey et al., 2017a) (Athey et al., 2017b) Some work has looked at housing markets with similar approaches. (Walls et al., 2018) (Royall, 2016) One recent work which uses a stochastic approach to identifying equilibria has tackled how changing consumer information affects the Tiebout hypothesis. Jehiel and Lamy (2018) The Athey et al. (2017a) specification uses a multidimensional prior for a value function that resembles a conjoint analysis utility function for consumer products. Jehiel and Lamy (2018) opts for a more complex period-over-period updating prior which embeds certain Markov assumptions in the utility function, somewhat akin to a hidden Markov Model (HMM).

3 Methodology

This paper models the survival function as an outcome variable of a demand process. In other words, individuals each have a value function $\max_j U(j)$ that determines their optimal neighborhood choice among neighborhoods $j \in \mathfrak{N}$. They choose to leave their current neighborhood when that value function is no longer optimized in their current neighborhood. This produces a neighborhood-level survival function $S_j(t)$ that aggregates the various lifetimes of all of the individuals in neighborhood j. Assume all of these decisions are according continuous time functions, which are observable in discrete time intervals (in our case, annually).

3.1 Hazard-Setting Game

Residents i of a given neighborhood j experience three forms of goods: amenities (public and private), housing goods (bedrooms, bathrooms, backyards, etc.), and access to local structural resources (e.g., employment opportunities, healthy food access, local bank branches, distance to central business district opportunities, etc.). As these values evolve after the period T in which the resident moves

in, individuals derive some utility according to an arbitrary function of the three goods and the local price level for each. This function is assumed to have the same functional form across individuals, but the form itself is unknown. Individual parameters of the value function, such as elasticity to various observable goods provided by the neighborhood, are heterogeneous across individuals.
Local policy decisions, changes in local employment markets, and city-level demographic shifts cause fluctuation in each of the goods over time. Thus, the goods value of a neighborhood is time-varying. Then, we can define a "local component" of the value function of the individual that is totally contributed by local features to the individual's neighborhood. If $U^{\mathcal{N}}_{ijT}$ corresponds same individual utility function but is entirely calculated via the *relative value* of local variables and parameters:

$$U^{\mathcal{N}}_{ijT} = U^{\mathcal{N}}_{ijT}(A_{jT}, \eta_{jT}, \Pi_{jT}, P_{jT}, d_{jT}, h_{jT}) \tag{2}$$

Terms correspond to *relative* values of goods: local amenities, employment access, and housing, and the costs thereof. By relative, we mean that the entire universe of neighborhoods is ranked relatively on these features and the resulting relative weights of each neighborhood is used to calculate the above function. Individuals leave when the utility of local goods is no longer justified by the cost. Note that individuals place varying weight on varying amenities and housing goods. This is captured by allowing varying utility assignments to $\{A_{jt}, \eta_{jt}, \Pi_{jt}\}$, which represent latent variables that aggregate the relative values of objective measures of local state variables.

Assumption 1.

Individuals know own-neighborhood latent aggregated state variables with sufficient accuracy.

With this in mind, we argue individuals, in a world with perfect mobility (an assumption which is relaxed later), will choose to move out when a benchmark, or "market-clearing" utility is breached. Continuing the language of relative weights, if each neighborhood has some total positive local-utility function to the individual of U^{+}_{ijt} that comes at some cost U^{-}_{ijt} - which includes the opportunity cost of the next best neighborhood, then the individual moves when

$$\bar{U}_{ij} = U^{+}_{ijt} - U^{-}_{ijt} < 0 \tag{3}$$

Using this threshold framework, I focus my analysis on learning the distribution of U^+_{ijt} and U^-_{ijt} across all individuals by neighborhood to create distributions, U^+_{jt} and U^-_{jt}. These are dimensionless "push" and "pull" factor distributions, the latent aggregated variables governing the decision to move out of a neighborhood. Individuals experience a push and pull factor drawn from these aggregated distributions in a given period, making their decision to churn based on these parameters. This approach leverages self-selection; residents select themselves to move out of a neighborhood based on these parameters, revealing their preference for the current neighborhood. A comparison in the literature is the Roy Framework used by labor economists to model the decision to enter the labor market and leave home production. (Heckman and Honore, 1990)

This problem lends itself well to a Hierarchical Bayesian model when aggregated to neighborhood-level distributions. We do not know the absolute survival function of a neighborhood. But, census data provides annual histograms of out-migration by income. By estimating a two-parameter distribution that describes out-migration frequency by income, we have an estimated $S_j(t)$ for that period. This is the outcome variable. We can also attempt to fit an error model to the margins of error that are associated with this histogram estimation procedure; this topic is left for a later date, but can provide added robustness to the machine learning procedure when included *a priori*.

In particular, if the aggregate income distribution of neighborhood j in period t is denoted Y_{jt}, but it is $Y_{j,t+1}$ in period $t+1$, then the out-migrant population $\Delta_{j,t+1} = Y_{j,t+1} - M_{jt} - Y_{jt}$ (with M_{jt} the income distribution of new migrants in this time-frame) had consumed at or below a clearing-price utility $\bar{U}_{ij}$ while the survivors consumed at or above this utility. This absolute binary is used to adjust a measure of the relative push and pull factor distributions of the neighborhood.

At a given time, the change distribution $\Delta_{j,t+1}$ indicates that $U^+_{ij} - U^-_{ij} < 0$ for $i \in \Delta_{j,t+1}$ and $U^+_{ij} - U^-_{ij} \geq 0$ for $i \notin \Delta_{j,t+1}$. We update our distribution of beliefs about $U^{\pm}_{jt}$ each period based on this data. What this exercise demonstrates is that any two-parameter distribution F with parameters that are levers for increasing or decreasing the likelihood of the outcome - correspond to $U^{\pm}_{ij}$ - can be used to model the heterogeneity in push/pull factor balance, $U_{jt} = U^+_{jt} - U^-_{jt}$. In particular, this suggests a beta-distributed model of behavior. Because we have a single outcome variable, the outcome distribution is singular, but the use of the beta implies that the hyperpriors for the two parameters, α and β, should mirror the maximum likelihood distributions of U^+_{jt} and U^-_{jt}, respectively. If every individual experiences the same utility function U_{ijt}, then the survival function would necessarily be an entire upheaval or an entirely stable neighborhood. If, however, individuals consume at variable rates of substitution, then we can construct a survival distribution across individuals. Let the probability that individual i leaves and that their value function $U_{jt} < 0$ conditional on income is given by the parameter θ_i. Then, individual hazards are

$$h_i(t) = f(\theta_{it}|y_i) = P\{\theta_{it}|\alpha_{jt}, \beta_{jt}\}P\{\alpha_{jt}, \beta_{jt}|y_i\} \tag{4}$$

where $f(\cdot)$ is the Bernoulli trial corresponding to the probability of moving out. Integrating across individuals,

$$\int_{\mathcal{N}_j} h_i(t)dt = S_j(t) = F(\Theta_{jt}|y_i) = P\{\Theta_{jt}|\alpha_{jt}, \beta_{jt}\}P\{\alpha_{jt}, \beta_{jt}|\Delta_{jt}\} \tag{5}$$

Keeping the structure of the hazard function simple is important for interpretation of the model output, so this paper will focus entirely on Beta-Binomial data generating processes (geometric and Weibull hazards were also considered, but dynamics are difficult to address with histogram data. See 3.6).

This structure implies that neighborhood-level push and pull factors are entirely identified if we know Δ_{jt}. The next sections conduct a posterior analysis of the learned structural parameters (α, β) and uses them to extract elasticities to various amenities, as well as a measure of residential inertia.

3.2 Two-Geography Case

Suppose residents are choosing between two different neighborhoods, North Philadelphia N and South Philadelphia S. Consider a simple logit model to begin. Residents choose to maximize a utility with respect to geography j:

$$u_{ijt} = \vec{\mu}A_{jt} - \gamma P_{jt} + \tau_t + f_j + (1/\rho)\epsilon_{ijt} \tag{6}$$

They derive this utility by comparing time-varying amenities A_{jt} of the neighborhood (as well as some fixed effects for neighborhood f_j which may correspond to persistent amenities over time and time-specific adjustments corresponding to exogenous area-level shocks τ_t) to the price they pay to live in the neighborhood P_{jt}. In the language of the above framework, $U^+_{ij} = \vec{\mu}A_{jt} + \tau_t + f_j$

and $U_{ij}^{-} = \gamma P_{jt}$. They have some vector of preferences for various amenities which defines their response to local conditions. For now, all local pull variables are considered to be a piece of A_{jt}, though these are disaggregated in subsequent sections. Allow ρ to represent the precision of unobserved preferences (i.e., $1/\rho$ is the logit scale parameter). We for now will assume no exogenous term τ_t, Individuals stay in the neighborhood with probabilities:

$$\mathcal{P}_N = \frac{\exp(\vec{\mu} A_{Nt}\rho - \gamma P_{Nt}\rho + f_N\rho)}{\exp(\vec{\mu} A_{Nt}\rho - \gamma P_{Nt}\rho + f_N\rho) + \exp(\vec{\beta} A_{St}\rho - \gamma P_{St}\rho + f_S\rho)} \tag{7}$$

$$\mathcal{P}_S = \frac{\exp(\vec{\mu} A_{St}\rho - \gamma P_{St}\rho + f_S\rho)}{\exp(\vec{\mu} A_{Nt}\rho - \gamma P_{Nt}\rho + f_N\rho) + \exp(\vec{\mu} A_{St}\rho - \gamma P_{St}\rho + f_S\rho)} \tag{8}$$

The resulting model suggests that a high value of ρ suggests high heterogeneity which will lead to high sensitivity to amenities, whereas a low value of ρ implies little impact of unobserved heterogeneity and thus a low sensitivity to changes in amenity. Individuals leave with these strictly positive probabilities. However, the model lacks identification of a vector $\vec{\beta}$ and therefore we cannot necessarily make any causal claims around utility.
In the framework we introduce above, however, we develop a beta-geometric probability of leaving the neighborhood. This model suggests that rather than the above logit, in a single period case we have

$$1 - \mathcal{P}_N(\Delta_N) = \frac{B(\alpha_N + 1, \Delta_N + \beta_N - 1)}{B(\alpha_N, \beta_N)} \tag{9}$$

$$1 - \mathcal{P}_S(\Delta_S) = \frac{B(\alpha_S + 1, \Delta_S + \beta_S - 1)}{B(\alpha_S, \beta_S)} \tag{10}$$

This system develops two descriptive parameters for each neighborhood, α_j and β_j. The expected probability of leaving a neighborhood is given by $E(\beta_j/\alpha_j)$. Why is this better? Because we can isolate from these two reduced form parameters distinct components of variation. The proportion of heterogeneity corresponding to individual heterogeneity in this model is:

$$\Sigma_j = \frac{Var(\beta_j/\alpha_j^3)}{E_j(\beta/\alpha^3)} \tag{11}$$

The remainder of the variation can be explained by structural parameters. Then, by estimating first these push and pull parameters and controlling for individual-level preferences, we can estimate the structural response to a set of time-varying amenities on these reduced form parameters. Consider the two city case. If we obtain a maximum likelihood estimate of the two regions, we obtain a heterogeneity-adjusted measure from:

$$\hat{\alpha}_N = \Sigma_N \alpha_N$$
$$\hat{\alpha}_S = \Sigma_S \alpha_S$$

and analogously for β.

The suggestive value function for this new model is then (scaling for money metric utility),

$$V_j = \log(\exp(\beta_j) - \exp(\alpha_j))$$

The simplicity of the value function is suggestive of a simple resulting elasticity estimation, obtained by ans application of the envelope theorem. I continue this after generalizing the two-geography case. In the rest of the paper, for convenience, we will drop the hat notation. Below, I discuss how the model developed here allows for the creation of a histogram-level analysis, a technique which enables Census data to generate valuable insights. I then discuss how we can use the reduced-form estimates to obtain a robust estimator of the elasticity of the neighborhood's geographic mobility to a particular amenity.

3.3 From Histogram to Parameters

Neighborhoods have a survival histogram (alternatively in the planning literature, a tenureship histogram) which describes the number of individuals who currently reside in the neighborhood and have been there for d years, for a range of values $0 < d \leq T$. There exists some mean duration $\bar{d}_j$ across residents in neighborhood j.

Individuals each have some latent calculated utility of residency, $U_{it} \sim \theta_{it}$ in each period. In other words, the beta-geometric is a projection of money-metric calculated utility from the real line to the unit interval. Values below 0.5 correspond to a negative utility, $U_{it}^{+} < U_{it}^{-}$. Individuals are more likely to leave the deeper negative their utility goes. Across the population, this θ_{it} is distributed according to some distribution that generates the unconditional histogram (binned density function) of survival $S_{j,t}$ for time frames

$\mathcal{D} = \{0 \leq d < d_1, d_1 \leq d < d_2, \ldots, d_{K-1} \leq d < d_K\}$.

Then, using the data generating process above, individuals have lower θ_{it} with a likelihood to a time- and neighborhood-specific pull factor β_{jt}, and a higher θ_{it} with the push factor α_{jt}. The push factors form a sort of "shadow price" of residency and the pull factors are in turn a utility of staying. These values should be distributed according to the population-level distribution of $\log(U_{jt}^{\mathcal{N}})^+$, or the logarithm of the positive piece of an individual's neighborhood "basket of goods" value function, and $\log(U_{jt}^{\mathcal{N}})^-$, respectively. The reason for the logarithm is to give a convenient form to the mean of the mean churn.

The mean probability that a resident chooses to leave the neighborhood corresponds to the balance between these two factors. The mean likelihood should be the expected value of utility, or $E((U_{jt}^{\mathcal{N}})^+ - (U_{jt}^{\mathcal{N}})^-)$. Without loss of generality, one can logarithmically transform utility so this likelihood is $E\log((U_{jt}^{\mathcal{N}})^+ - (U_{jt}^{\mathcal{N}})^-)$. Concavity of the logarithm implies this quantity is strictly above $\log(E(U_{jt}^{\mathcal{N}})^+) - \log(E(U_{jt}^{\mathcal{N}})^-) = \frac{\log E(U_{jt}^{\mathcal{N}})^+}{\log E(U_{jt}^{\mathcal{N}})^-}$. Then, a reasonable expression of the mean likelihood to leave the neighborhood comprises a ratio of the push and pull factor, defined to abstract away the logarithmic algebra: $E(\theta_{it}) = \frac{\beta_{jt}}{\alpha_{jt}}$. In other words, the mean of the empirical histogram $S_{j,t}$ should provide an estimator of the ratio of the push and pull factor of the neighborhood. A higher mean implies either a higher pull or lower pull factor - which stands to reason.

The model uses this logic to generate a hierarchical model using a histogram of survivals rather than explicit microdata. Provided an uninformative prior for α_{jt} and β_{jt} which assumes uniform positive and negative utility across the population, a Monte Carlo Markov Chain (MCMC) estimator then calculates the theoretical likelihood of obtaining $S_{j,t}$ given these distributions for α_{jt} and β_{jt}. Note that the maximum likelihood estimate generally aligns with a method-of-moments approach, but the use of histogram data prevents reliable estimation of a variance measure. Therefore, testing the likelihood of the full distribution $S_{j,t}$ is a more stable empirical procedure for estimating α_{jt} and β_{jt}. In other words, rather than estimating moments, the maximum-likelihood estimate (which requires MCMC estimation for this hierarchical setting) maximizes $N_{jt}P\{d_{k-1} \leq d < d_k\} = S_{j,t}(k), k \in \{1, \ldots, K\}$ - the theoretical likelihood of the histogram itself.

Then, in summary, our procedure is:

1. For each neighborhood, calculate the out-migration population histogram

for each year, Δ_{jt}, where the histogram is along an axis of resident duration time described as $\mathcal{D}$ above.

2. Use MCMC estimation to find the posterior likelihood for α_j and β_j based on the sequence histogramΔ_{jT}. The mean duration β_j/α_j corresponds to the mean $E_{\mathcal{D}}(\Delta_{jt})$. Label these values α_{jT} and β_{jT}.

3. Repeat the procedure for each histogram in the series, $\{\Delta_{j1}, \cdots, \Delta_{j,T-1}\}$. Label these values α_{jt} and β_{jt} by time period.

4. Calculate the estimated standard error of β/α^3 in each year across neighborhoods, as well as the true $E(\beta/\alpha^3)$ given the MCMC likelihood. The ratio between these gives the distribution of Σ_{jt}.

5. Perform a correction which removes the portion of β_{jt} and α_{jt} which is not explained by variation in composition of neighborhoods so that the remaining values are precisely those attributable to structural differences between neighborhoods, rather than individual preferences.

With this calibration technique in place, I now describe how we can obtain identifiable elasticities.

3.4 Elasticity Estimation

The benefit of a hierarchical Bayesian model is the ability to minimize shape assumptions of the utility model. Because we use a histogrammed dataset, having the model be responsive to sampling variation is extremely important. Begin with the push-side. Propose the following form for the latent push factor, α_{jt}, based on a mixed proportional hazards-type model:

$$\alpha_{jt} = H(a_j + \gamma_1 P^-_{jt} + \gamma_2 \sigma_{jt} + \tau_t t + \epsilon^{(\alpha)}_{jt}) \tag{12}$$

where the components correspond to: a latent baseline (pulled from some uninformative prior, say a uniform distribution), a simple linear term in relative housing prices in this neighborhood, the same in the deviation of the local income distribution, a time-specific term to capture taste shocks, and an idiosyncratic shock. The negative superscript on the housing price latent variable refers is the difference between objective market assessments of a hedonic house price and actual annual sale prices from a property assessment agency. This residual corresponds to the component of prices of the housing good attributable to local price levels (plus an idiosyncratic error) in that year. $H(\cdot)$ refers to a transformation (like an inverse normal transformation) which makes the expression into a digestible parameter for the mixture model. There is no economic reason to leave a neighborhood other than prices and a potential volatility in pull factors. This latter factor may seem redundant with the fact that the pull factor latent variable will shrink as pull factors like amenities and income are less readily available, but as individuals often tend to anticipate such future shocks to an economy, these terms allows speculative shocks to grip a neighborhood. Notice that, much like in Athey et al. (2017b), most of these parameters can be learned in a pooled fashion, so that observations across neighborhoods share statistical power.

The pull factors of the economy are local income and amenities. By local income, we mean the component of income which is determined by local employment access and access to vacancies. To determine this, we craft a new variable η_j which defines local employment access. η_j may alternatively be seen as a ratio between the current level of employment in the neighborhood versus full employment. It is hard to measure, of course, but might be proxied by known variables. Consider

$$\bar{Y}_{jt} = Y_{jt} - \frac{\sum_j Y_{jt}}{\sum_j N_{jt}} =: Y_{jt} - \tilde{Y}_{jt} \tag{13}$$

The difference between the income *distribution* in a given neighborhood and the distribution of the sample mean of the income distribution, across all sampled neighborhoods. In an unbiased sample of neighborhoods, this difference is the part of the local income distribution which deviates from simple expectations. Theoretically, this difference should be explained entirely by local spatial effects and heterogeneity effects (composition). In other words, people in a neighborhood with far higher income than the income of a city overall are either more likely to have higher income or receive distinct advantages from the neighborhood (more likely both). Further, such an income distribution in order to remain high-income as people churn out must consistently have high-income people - this is another way of saying there is low relative standard deviation in incomes relative to the overall distribution of income in the sample. Estimate these two effects as:

$$\bar{Y}_{jt} = \eta_{jt} + \omega_1 \frac{\sigma_j}{\sigma} + \omega_2 \frac{E(Y_{jt})}{\tilde{Y}_{jt}} + \omega_3 t \tag{14}$$

The above equation implies the relative standing of neighborhood j is a function of the local employment access variable plus a function of its relative variability (more variable neighborhoods will naturally possess higher means for strictly positive income) and its relative mean income. The terms with coefficients ω_k should capture the compositional effects, again pooling the ω_k across neighborhoods. This measure of η_j can be improved further in a broader structural model of the individual. One can imagine a duration dependent measure of η_j that mirrors how skill loss is modeled by labor economists studying unemployment. (Mueller et al., 2018) In particular, the longer an individual spends in an area which appears to under-employ them, the more their skills deteriorate and thus the stronger the impact on their income. This creates a mechanism for hazard rates in poverty-trap regions.

Another way to express this would be that there are three components to variations in sample means; compositional effects, heterogeneity effects, and fixed effects. Compositional differences should be explained by the ratio of means, while the degree of heterogeneity around compositional differences is tested for by the first term. The fixed effect remains, which describes a neighborhood's inherent employment access level bonus. Higher access means lower wage loss from travel time, lower chance of underemployment due to higher access to a range of vacancies, and thus likely higher wages. Not every neighborhood will have strong compositional effects or fixed effects, but most should have one or the other.

Once we calculate η_j, the pull-side regression comes together cleanly.

$$\beta_{jt} = H(b_j + \zeta_1 \eta_j + \zeta_2 A_{jt} + \zeta_3 \Pi^+_{jt} + \delta_t t + \epsilon^{(\beta)}_{jt}) \tag{15}$$

The same definitions for individual components is used as above. The positive aspect of the price variable refers to market value based on a hedonic valuation of assessed characteristics of the house (bedrooms, bathrooms, and physical features). This represents the "goods" being consumed when buying housing in the area. A_{jt} refers to a measure of local amenity level. Like the local employment component, this variable may require some clever measurement techniques. Unlike η_j, however, this will be more rooted in the data available than an econometric formulation.

One may ask how assignment to either estimator is chosen, as push factors should theoretically correspond to negative pull factors. Consider a subset of neighborhoods with some fixed, equivalent α_{jt} but different values of β_{jt}. Then, the true pull factors should account for the differences in β_{jt} given the same α_{jt} across neighborhoods. In this way, the sufficient statistics become identifiable quasi-experimental variation in one another. Then, there is a rejection standard for these posterior models, as variables can be uncorrelated with their assigned latent component once conditioning on co-requisite latent variables is taken into account. Then, consider the following simple estimation technique:

1. Match neighborhoods into groups with similar values of α_{jt} using a clustering technique such as k-means or Principal Component Analysis.
2. Conduct the posterior analysis of β_{jt} in these neighborhoods.
3. Repeat for clusters on β_{jt} to obtain estimates for a posterior decomposition of α_{jt}.

This structural estimation procedure is in a sense a matching method which ensures that neighborhoods only vary with regards to one parameter. We can now estimate the elasticity of churn with respect to a particular neighborhood attribute. Consider the quantity

$$\epsilon_{it}^{\chi} = 1/\sum_{j=1}^{\text{card}(\mathcal{N})} \frac{\partial h_j(t)}{\partial \chi} \tag{16}$$

where χ is some explanatory variable. Then, the elasticity is a distribution measure.

$$\frac{\partial h_j(t)}{\partial \chi} = \sum_i \frac{\partial \theta_{it}}{\partial \chi} \tag{17}$$

$$P(\frac{\partial h_j(t)}{\partial \chi}) = \frac{\partial P(\theta_{it})}{\partial \chi} \tag{18}$$

It is considerably easier to estimate the logarithmic case of these derivatives, e.g., to estimate the partial of the logarithm of the hazard rate, for reasons which will be apparent below.

$$\frac{\partial \log P(h_j(t))}{\partial \chi} = \frac{\partial \log P(\theta_{it})}{\partial \alpha_{jt}} \frac{\partial \alpha_{jt}}{\partial \chi} + \frac{\partial \log P(\theta_{it})}{\partial \beta_{jt}} \frac{\partial \beta_{jt}}{\partial \chi} \tag{19}$$

$$= \frac{\partial \log P(\theta_{it})}{\partial \alpha_{jt}} \gamma_\chi + \frac{\partial \log P(\theta_{it})}{\partial \beta_{jt}} \zeta_\chi \tag{20}$$

$$= \frac{\partial L(\Theta_{jt})}{\partial \alpha_{jt}} \gamma_\chi + \frac{\partial L(\Theta_{jt})}{\partial \beta_{jt}} \zeta_\chi \tag{21}$$

In other words, the elasticity of churn to a particular variable is the inner product of the score of the log-likelihood on the parameter space Θ_{jt} and the coefficients of that particular state variable in the model above. Note that the log-likelihood of the beta-geometric setting is defined as

$$\log L = L(\Theta_{jt}) = \sum_{i=1}^{n_j} \log B(\alpha_{jt} + 1, x_i + \beta_{jt}) - n_j \log B(\alpha_{jt}, \beta_{jt}) \tag{22}$$

where n_j is the population of the neighborhood under consideration. This is a convenient characterization, as it means we can completely characterize the structural response of a neighborhood to a particular shift in state variables through the model above and the log-likelihood. In other words, we can derive a secondary measure of demand sensitivity to a particular element of the "basket of goods" across a neighborhood based on churn data in this parametrization.

3.5 Inertia

Not every piece of the mobility equation is entirely measurable. Further, mobility is not necessarily free. and this is often the critical failing of a simple Tiebout process. Individuals may experience a hefty push factor from their current neighborhood but be kept in place by insurmountable search costs. Others may choose not to leave due to unique and un-measured community effects. The above model should therefore consistently underestimate the push and pull factors of the neighborhood using empirics - even highly specified empirics like the local component of income discussion above. Then, I propose training a second model with a specific inertia parameter.

A principled analysis of when the *a posteriori* break-down of Beta-Bernoulli model results fails beyond omitted variables produces the following four explanations:

1. The individual believes they will not find a better option elsewhere.

2. Community effects significantly alter the economic decision to move.

3. Some nonstationary, exogenous force shocks the specification in a manner that does not neatly fit one of the above parameters.

4. The individual cannot move due to other physical, legal, or un-captured political reason.

On nonstationarity: in the current model, this nonstationarity would likely be absorbed by a modeler ignorant of their source by the time fixed effects.
On beliefs: it's possible individual beliefs are responsible for the remaining discrepancy between model and fact. Perhaps a planning fallacy or present-biased mentality causes discounting of the utility benefits provided by other neighborhoods. This paper largely eschews a dynamical approach for a simpler analysis, however, making beliefs estimation a subject for a later model.
Community effects are difficult to estimate in a truly quantitative manner. In this setting, we can control for some degree of community effects by including measures of racial evenness (the degree to which racial groups are evenly dispersed throughout a neighborhood of nearby census tracts) and income bracket exposure (probability that individuals in different income brackets live in close proximity, assuming a uniform meeting probability with everyone in the neighborhood). Other sources of community, like a common ethnic identity or presence of strong institutions like a church or Catholic school, need to be assessed case-by-case. I test some examples ex post.
Introduce, therefore, an inertia to the posterior analysis. A moving-out penalty must be strong enough not just to drive the utility of moving to 0 (or even crafting a disutility to moving out), and as a result the hazard rate should drop to either 0 or an arbitrarily low hazard rate floor. The generating process is still the same as above, but now we add this penalty f_j:

$$\tilde{h}_i(t) = f(\theta_j | \alpha_{jt}, \tilde{\beta}_{jt}) \tag{23}$$

$$\tilde{\beta}_{jt} = H(\min\{f_j(b_j + \zeta_1 \eta_j + \zeta_2 A_{jt} + \zeta_3 \Pi^+_{jt} + \delta_t t + \epsilon^{(\beta)}_{jt}), b_j\}) \tag{24}$$

Inertia is interpreted as a forced stay; it does not allow one to leave if the pull factors fall below a neighborhood-specific critical threshold. Fixed effects actively lower the push factor shadow price of staying in the neighborhood in the current period. For individuals whose income level is low enough that they are likely to have high elasticities to changes in push factors, this term should have a large effect size. A higher fixed cost can drive the parameter for action, this α parameter, all the way down to the base case a_j (meaning the minimum churn parameter for this neighborhood - minimal movement out of the neighborhood year over year).

This is an important deviation from the originating dynamics which pegged this pull-side hazard parameter to the positive part of a utility function. Such a fixed cost may enter as a lower income in a larger period-by-period income strategy in a utility setting with full dynamics, but in a probabilistic setting considering churn in period t, the fixed cost is presented each period (it is not sunk but an active consideration in each period).

Once f_j is estimated, we can compare it to other indicators of a neighborhood's health and determine whether it varies in relation to other economic phenomena like neighborhood blight, credit tightness, or land use restrictions. This may suggest policy avenues for addressing areas with particularly low mobility and high inertia.

3.6 Dynamics

The current model does not allow for duration dependence or dynamics in the hierarchical beta-binomial model. The key effect is due to composition, or heterogeneity in the underlying population. Such a structure rules out the possibility of dynamic, Markovian hazards. With more complete individual-level data on movements, this dynamic setting would likely be more realistic and more possible on the machine learning side. Dynamics can be investigated in the posterior space, however.

4 Data

This project leverages publicly available data on Philadelphia's neighborhoods. Demographic and housing variables are derived from the American Community Survey. Variables are available from 2010 to 2016 at the census tract level. There are 384 census tracts in Philadelphia, with around 20 tracts having 0 or negligible population in most years. This could be due to these tracts contained large parks or transient populations, or simply due to accounting or measurement conventions. There is a tradeoff to using publicly available ACS data. To

get this level of granularity requires use of 5-year averages. This has the added benefit of smoothing out nonstationarities and discontinuities in variables due to uncaptured chance events. But, we lose information in the process of averaging. Recognizing this, I still make the choice knowing that with fuller microdata the analysis can be replicated with proper controls on chance nonstationarities. I follow the lexical convention of allowing the 5-year average datapoint to represent the middle year of the average. For example, 2011-2015 represents the 2013 year.
Housing microdata samples are obtained using public records data from various Philadelphia city departments. The official appraisals of a random sample of homes across census tracts is afforded by the Philadelphia Office of Public Assessment dating back to 1935, and up to the current year. These appraisals include a model for a hedonic price and a breakdown of housing characteristics including bedrooms, bathrooms, lot sizes, and other physical features. They also include a sale price which represents the hedonic value plus value-add from local amenities and other sources. The property assessment data includes group quarters, apartments, and other rental properties.

The homeowner population in Philadelphia is static or decreasing as the housing stock dates back in many regions of the city to the 1940s, making gentrification and disinvestment as forces primarily renter-focused. While income of individuals in old housing stock is stable, shocks to homeownership tend to imply large vacancies rather than gentrification. Many of these blighted neighborhoods, with large vacant blocks (ex. West Kensington) have not since been gentrified.

This data also provides features on housing characteristics (such as the number of bedrooms and other livability variables). Additionally, permitting and zoning data for all of these neighborhoods are available as well, where new construction permits are an important indicator of new development and gentrification.

Amenity-related data is also utilized. This data is drawn from the Google Maps API and Google Places API.

As is conventional in the economics literature, gentrification is primarily defined as a change in income. Figure (1) shows the distribution of changes in mean neighborhood income over the sample period.

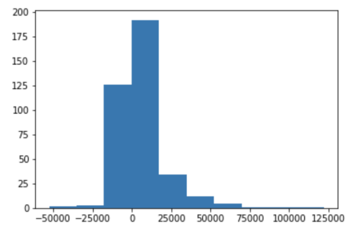


Figure 1: Changes in neighborhood mean income over the sample period.

Survival rates for individuals of low, middle, and high income using the income bracket cutoffs defined in ACS table B19001 are shown in Figure (2).

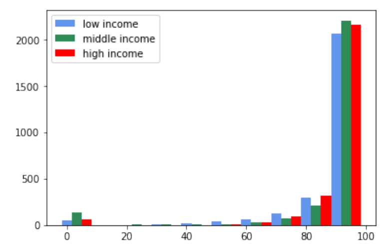


Figure 2: Histogram mapping survival rates by income bracket. (%)

Generally, the survival rates seem relatively homogeneous across the three income brackets, without strong significant differences. However, this is aggregated across neighborhoods. In specific neighborhoods, these numbers diverge widely, suggesting that local heterogeneity drives survival but that overall local heterogeneity is zero-sum - to be expected.

5 Calibration

The entire model is calibrated using Markov Chain Monte Carlo methods. The MCMC methodology was chosen over variational inference due to the use of aggregation and a need for greater control over potential bias in the model (VI tends to propagate bias in a manner that is more difficult to measure than in Hamiltonian Monte Carlo). (Betancourt, 2017) First, I report estimation of the key observables used in posterior analysis. Then, I report the machine learning model results for beta-binomial latent parameters and posterior analysis with observables. Finally, I report the results of an inertia-augmented posterior results.

5.1 Estimating Observables: Local Employment

Estimating observables of the model requires estimation of η_{jt} as in equation (13). A plot of the empirical distribution of the known regressors is shown in the figure.

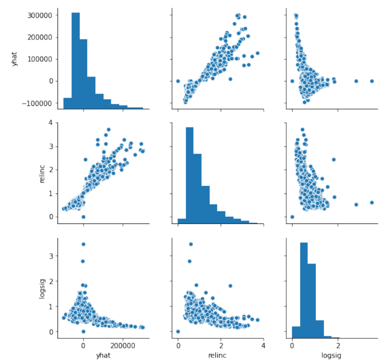


Figure 3: Matrix plot of key regressions in (13). From top to bottom/right to left: y-hat, relative income ratio, and log of the ratio of standard deviations.

The variables $\bar{Y}_{jt}$ and 'relinc', which refers to $\tilde{Y}_{jt}$ are both somewhat reminiscent of a gamma distribution, with some evidence of hyperdispersion. Meanwhile, the logarithm of the standard deviation ratio in (13) - which is presented in a logarithmic form to highlight some of the grain in the relationship - is most informative at higher levels of the level of between-neighborhood heterogeneity, with a significant amount of jitter around 0 (implying neighborhoods around the mean have some fairly variable shapes). We can expect the model to have some difficulty capturing this and experimented with binning the relative sigma values to make the regression more informative.

Before directly producing the model described in (13), we fit a model without an annualized ω_3 and allow for a different year over year value of the coefficients ω_1 and ω_2. This determines whether the effect of time is a fixed one exogenous to the given regressors or whether it is in fact a change in the regressor effect size. The results of an MCMC run are shown in Figure (4), with average effective n of about 1000. The figure presents a few important conclusions; we treat them in kind. First, that the effect estimated mean of the relative sigma coefficient is fairly strongly negative; in otherwise a higher relative variability indicates a lower local income distribution mean overall. This seems to indicate variability in this Philadelphia sample is biased more strongly left of the mean - which makes sense given Philadelphia has strongly spatially segregated poverty.

On a related note, then, is η tends to have a lower value - in other words, local structural access is more likely to pose a disadvantage than an advantage. This is a complex variable to decompose, and can range in interpretation from travel time to access to government support services in finding new work.

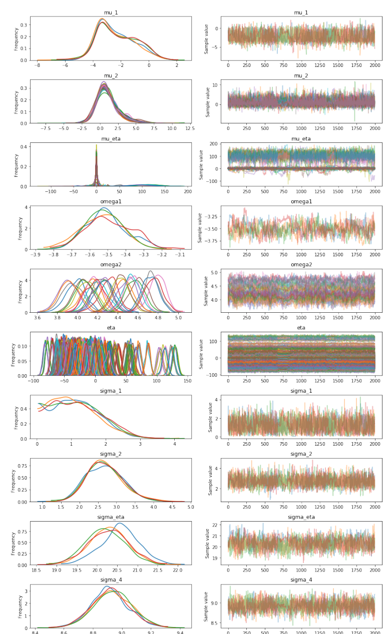


Figure 4: MCMC run with time-based coefficients.

I compare these results to those where (13) is reproduced directly, shown in Figure (5). The new model has a widely available information criterion (WAIC) of 19629, compared to one of 19928 in the previous setting, indicating a potentially better sampling of the posterior space of η. However, we also have a higher standard error on the WAIC in this new model, which suggests the fit is less consistently good throughout the sample.

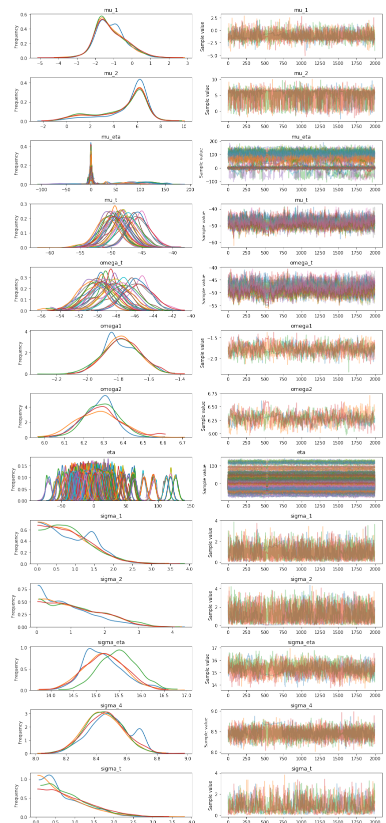


Figure 5: MCMC run with a separate time term.

In the new model, we see that the general distribution of ω_1 and the terms η_{jt} is fairly robust to the shift in model. ω_2 is more cleanly positive, but has a long left tail - incomes have a strong effect but interestingly there is probability in the tails which indicates that in some neighborhoods the role of income variability or local effect may take over in explaining local spatial differences. The graph also shows a series of normals which generally monotonically shift right with time, as one may expect from a period of growth.

The mean η_{jt} are plotted against the tract-level mean churn times for middle income individuals in Figure (6). The color-bar shows the value of $\hat{Y}_{jt}$ in 2010. Clearly, higher income neighborhoods have generally higher employment access. Lower η_{jt} does not guarantee a lower level of $\hat{Y}_{jt}$, indicating that some of these low-lying neighborhoods are likely income-variable. High churn rates tend to happen in the neighborhoods without significant locational advantages or disadvantages - closer to the mean. Why these churn rates don't happen uniformly given local income differentials and η_{jt} is the subject of our model, which studies it in relation to estimated amenities and fixed costs of moving.

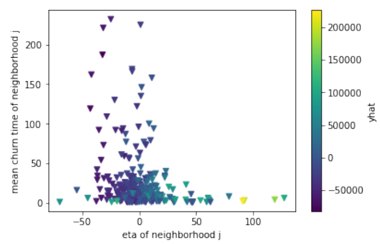


Figure 6: Scatterplot of local employment access η_j and the mean churn time for middle income residents in a given neighborhood in 2010.

5.2 Estimating Observables: Local Amenities

The next step in model estimation is defining how we can measure or proxy local amenities, A_{jt}. To do this, we draw on the Google Maps API and look at a selected vector of possible key amenities for each home observed in the Property Assessments database. Then, the distance and rating for each observed local amenity are factored in to come up with a local amenity score. As the key consideration in any amenity model tends to be the relative scoring, argue that a sufficiently inclusive model which captures these relative effects should capture this.

Amenity scoring can also have an affordability component; we can define this by again using the Google Maps API and looking at features like local restaurants and accessing price ratings of these features. By aggregating this information, we can score affordability of amenities in addition to their core presence score from above.

The amenities we look at are:

1. Parks and local green space.
2. Restaurants and cafés.
3. Grocery stores.
4. Corner stores and gas stations.
5. "Things to do" including cinemas, bowling alleys, and other entertainment/leisure venues.

6. Local crime rates.

7. Ratings for local schools.

8. Vacancies, measures of local housing quality, and presence of dilapidated housing.

In reality, a good amenity model will consider that individuals have potentially income-specific preferences and utility for various neighborhood amenities. We test the need for this in our model by allowing γ_2 and A_{jt} in (15) to exist as vectors. This allows variable estimation of relative utility. Note that a principled derivation of the utility function of an individual in neighborhood j is fully retrievable once these distributions are known, even in this hazard-based setting. Then, the key determination for including a vector of weights for various types of amenities is the numerical cleanliness of doing so.

We can craft a test to ensure our relative amenity weights are informative, furthermore, by testing whether they can explain a spatial fixed effect difference in market value and sale price of a set of homes. Property assessment data provides both of these factors. Market value in this dataset reflects a hedonic value purely derived by a model based on observables such as lot size, number of bedrooms, number of bathrooms, environmental hazards like asbestos, and other physical factors conventionally used as controls. The sale price of a house should be a function of this valuation, which acts as a sort of stationary anchoring of the sale price. Sale prices then vary additionally according to local amenity quality in a given year and the local community variables which are unobserved. Then, a good measure of amenities should behave well in the linear model:

$$P_{ijt} = val_{ijt} + \vec{\omega}\vec{A_{jt}} + \mu_t + \epsilon_{ijt} \tag{25}$$

Price of house i at time t in tract j has components equivalent to a hedonic value plus amenities plus a time fixed effect (with idiosyncratic error). See Figure (7) for a clear linear relationship between value and sales price exhibited in the Office of Property Assessment Data for 2010 through 2016. While property assessment data has historically been biased, we rely on the simple hedonic definition of value applied in the dataset here as a transparent and reliable one. With this in mind, we see a case for (25).

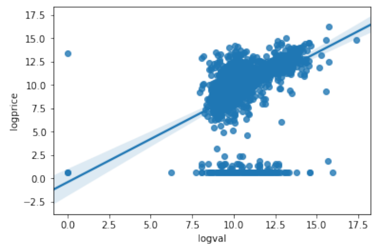


Figure 7: Simple bivariate linear case showing the logarithm of price plotted against the logarithm of the OPA market value calculation.

We test to see whether the Google Maps information we collected fits this expectation in Figure (8). These results are generally indeed satisfying. Distances which are significant indicate a growing distance from urban unfavorables (like a gas station) may increase price, while favorables (like a train station) have a large negative impact on prices. Higher rated amenities across the board *tend* to have positive impacts on price, with groceries and gasoline access being paramount to Philadelphia residents.

	coef	std err	z	P>|z|	[0.025	0.975]
groc_dist	0.0547	0.294	0.186	0.852	-0.521	0.631
gas_dist	0.5117	0.237	2.159	0.031	0.047	0.976
rest_dist	-0.7692	0.385	-1.999	0.046	-1.523	-0.015
night_dist	-0.0508	0.134	-0.378	0.705	-0.314	0.212
train_dist	-3.2737	1.603	-2.042	0.041	-6.415	-0.132
bus_dist	1.9524	1.495	1.306	0.191	-0.977	4.882
groc_rate	0.9813	0.402	2.442	0.015	0.194	1.769
gas_rate	1.1425	0.313	3.649	0.000	0.529	1.756
rest_rate	0.3448	0.333	1.034	0.301	-0.308	0.998
night_rate	0.4124	0.386	1.068	0.285	-0.344	1.169
park_dist	0.0059	0.044	0.134	0.893	-0.081	0.093
office_dist	-0.0287	0.061	-0.471	0.638	-0.148	0.091
park_rate	0.0591	0.448	0.132	0.895	-0.819	0.938
price_std	4.029e-07	2.99e-08	13.467	0.000	3.44e-07	4.62e-07
value	4.132e-07	3.56e-08	11.609	0.000	3.43e-07	4.83e-07

Figure 8: Robust linear model for sale price of a home based on its hedonic value and various measures of amenities.

Of course, other specifications were tested. Importantly, one specification which threw every variable which was not highly collinear with ‘value‘ in the dataset into the modeler and allowed backwards selection of covariates was selected; the final model was fairly parsimonious and smaller than the above, with the only notable addition being mean travel time to work.

Other measures of amenities like distance to a central business district (CBD) or nearby commercial hub also have value; but by disaggregating specific amenities in this way, one can ensure that a sufficient amount of heterogeneity of preferences is being exhibited to capture a larger variety of neighborhood demand. Because this work is particularly aiming to analyze blighted neighborhoods, assuming access to known commercial corridors or the CBD should be the main amenity considered by those still in the neighborhood is likely flawed; it's merely one piece of a true internal calculus. We include CBD via the "office_dist" variable in the regression - one can see it is not particularly robust in the face of the other regressors.

5.3 Beta-Binomial and Posterior Analysis

With all observables ready to use, I run the beta-binomial learning algorithm and run the described posterior analysis using observable variables. The coefficient posterior distributions are reported in Appendix B. These coefficients represent a measure of the elasticity of the likelihood of moving out to a positive change in the variable in question. Most of these sensitivities are significant.
Before continuing with posterior analysis, I briefly discuss the results of the beta-binomial fit. The distribution in 2016 in tract 157 of the latent parameters of the beta is shown in Figure (9). Clearly, this tract has a strong bias towards churning out of the tract. Note that while the horizontal axis is all negative, this does not translate directly to U_{ijt} being negative for all individuals; this analysis requires drawing a parameter θ_{it} from the distribution.

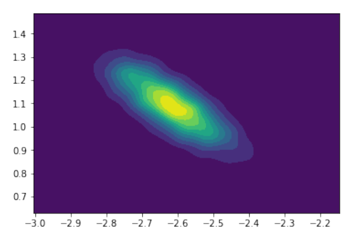


Figure 9: Horizontal variation is variation in $\beta_{jt} - \alpha jt$, vertical represents variation in $\beta_{jt} + \alpha jt$. Joint distribution of β_{jt} and αjt in tract 157 in 2016.

I also map the pull factors of each neighborhood at the end of the observed period over a map of Philadelphia. The distribution is mapped in Figure (10).

The maps shows a few expected trends. For example, high pull-factor areas include key commercial corridors like South Philadelphia and the Central Business District as well as wealthier districts on the outskirts of the city. Areas

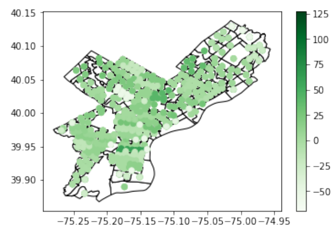
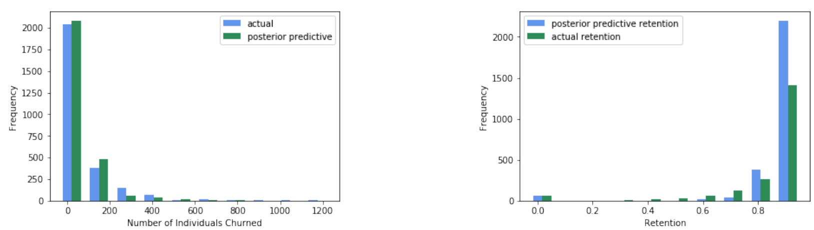


Figure 10: Map demonstrating values of the latent employment variable across Philadelphia census tracts.



(a) Histogram counting number of tract-year paired observations (e.g. an observation is tract j in year t) with specified population out-migrating.

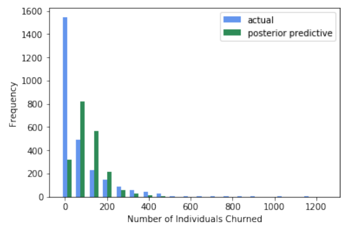
(b) Retention rates (e.g., proportion of population which survived) histogram with observations corresponding to tract-year.

Figure 11: Model calibration results.

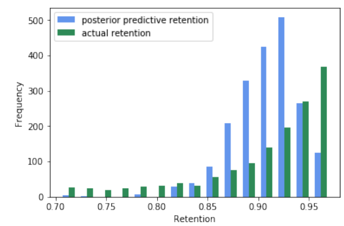
with persistent poverty issues, Oxford and 12th for example, are not as strong on pull factor. Newly gentrified Fishtown shows a very strong pull factor. This correspondence with fact is encouraging.

Returning to posterior analysis, more important than the estimated coefficients would be the reported fit. We analyze the fit using two histograms; the first, the histogram for raw number of individuals churned by neighborhood. The second, the histogram of retention rates by neighborhood ($\frac{1}{n_j}\sum_{i=1}^{n_j}(1-\theta_i)$). The first histogram, (11a), appears near perfect; there is a slight overestimation of the lowest two bins (this could be due to the need for a "zero-inflated" setting which captures the probability of a tract with 0 population, e.g., the probability of a measurement error). The second, (11b), tells a more nuanced story. The posterior semi-linear model dramatically underestimates very high retention rates.

To test whether the underprediction is the result of measurement error (0-population tracts) skewing the model prediction, we also test a zero-inflated model that explicitly assigns a non-zero probability to measurement errors in the



(a) Zero-inflated model performance on out-migration numbers.



(b) Zero-inflated model performance on non-zero tract retention.

Figure 12: Zero-inflated calibration results.

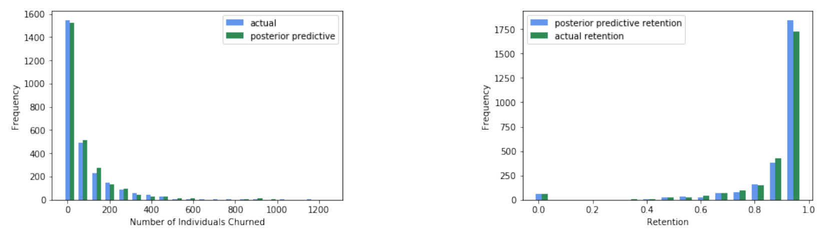
ACS. In practice, this means a a Bernoulli trial whose success event corresponds to the population of a tract being 0 is assigned for each neighborhood before the posterior model is run. With this new setting, we replicate the figures above on the datapoints which do not default to 0 (e.g., the subset of tract observations where $N_j > 0$).

Figure (12) clearly demonstrates that there is still significant under-prediction of neighborhoods with low out-migration and overprediction of high out-migration. We now investigate whether an inertia can resolve this contention.

5.4 Inertia Posterior Analysis

This model takes the extended equation (23) and implements it. The prior on inertia is a positive half-normal that spikes at 1 (because an inertia of 1 means the original specification is valid). If the posterior analysis is correct, the effect of including f_j on shape of the existing retention rate of neighborhoods should have either minimal or a slightly positive impact, raising the number of people predicted to stay. Potentially f_j should vary significantly with key covariates of the costs of mobility, such as local renter proportions, mortgage rates, house prices, and even local income effects measured by η_j. We test both of these as possible posterior predictive checks in addition to the usual histogram comparisons. Critically, note that inertia should primarily effect churn in neighborhoods with a sizable low income population as these are neighborhoods with a likely lower elasticity to changes in local amenities and other pull factors.

If mortgage and renter variables do not bear a critical effect on the inertia, yet f_j tends to be nonzero across neighborhoods, we argue that low-income community effects are the likely driver (see Section 3.5). In other words, a non-fiscal mobility inertia exists in this case.



(a) Raw population churned across all observations from model and observation.

(b) Retention rates across all observations from model and observation.

Figure 13: Model calibration results.

The posterior predictive histograms are shown in Figure (13). Clearly, the fit has resolved the issue of underpredicting low-churn tracts (in fact, there is a slight overprediction), and is nearly perfect across all bins of both the churn and retention histograms in a visual test. While a priori, the model was going to be tested with a zero-inflated fit, these results indicate such an approach might be mis-specified. Zero-population tracts were dropped. The new model's WAIC has improved approximately 3-fold from 314,352 to 115,585.

The distribution of the inertia, which recall is essentially a proportion of $\tilde{\beta}_{jt}$, is illustrated in the figure (14). The value, 'pf' (for "proportion f") exhibits bimodal behavior, with a smaller concentration of tracts having high (near-zero) inertia and a large section of tracts concentrated around the mean at 33% of the pull factor latent variable. Considering that Philadelphia is indeed the country's largest poor city, this first mode makes sense. The tail seems to indicate the richest neighborhoods which have little restriction to mobility and highly non-local communities.s

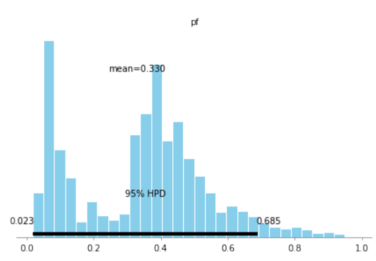
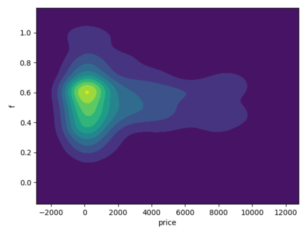
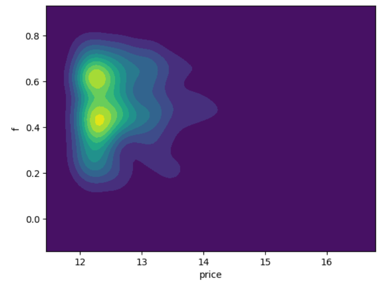


Figure 14: The estimated posterior distribution of the inertia term.

Now we test whether this estimator indeed can be explained as a inertia estimator. First, we look at the conditional distributions of inertia with respect to the price level of a neighborhood by comparing its distributions in census tracts in the top and bottom quartiles of local house prices.



(a) Inertia estimator in neighborhoods in the bottom quartile for home prices.



(b) Inertia estimator in neighborhoods in the top quartile for home prices.

Figure 15: Conditional distribution of fixed costs given local home prices.

Figure (15) does precisely this. Notice that in the higher income tracts, a second mode emerges above 0.5, suggesting highly mobile neighborhoods are more common in the high-price set. is being contributed by high-price neighborhoods.

Then, wealthier neighborhoods with higher local prices exhibit a tendency towards mobility based on the variable 'f' alone. This is suggestive that inertia also contains an income-specific fixed cost effect. With this correlation-based analysis, the hypothesis is now tested by using a suggestive linear model. Fixed costs should be a deterministic function of whether individual residents are homeowners, homeowners' mortgage paperwork, job search time. Job search time is generally a function of job finding likelihood (said at risk of upsetting the labor literature - this statement is of course a simplification of a complex search-and-matching process). The local job market's relative value, η will serve as proxy for an individual's likelihood of finding a future job. The two are fundamentally linked values, where job-finding likelihoods are both in reality and perceptions set according to localized job market accessibility, as argued by (Ioannides and Datcher Loury, 2004) and (Van der Klaauw and Van Ours, 2003). With this set of covariates, the following linear model is formed:

$$f_j = \beta_0 + \beta_1 \text{mean\_mortgage\_rate}_{jt} + \beta_2 \eta_j + u_t + \epsilon_{jt} \tag{26}$$

Some conclusions from the model are worth highlighting. First, higher rate of mortgage coincides with higher local inertia. Rentership coincides with lower inertia. Higher local employment opportunity and inertia are somewhat correlated, suggesting more predictably mobile neighborhoods have more opportunities as well. All of these values are significant in the measurement of f_j, even when robust standard errors with local- and time- fixed effects are used.

Variable	Estimator	p-value
rentership	0.118	0.009
mortgage	-0.245	0.001
η_j	0.001	0.065

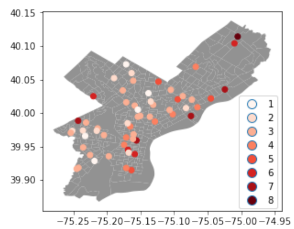
Table 1: Results of the regression in (26).

This analysis indicates that a source of local retention dynamics is that geographic mobility is significantly hampered by the fixed cost of leaving. Further, these fixed costs are tied up with low-income neighborhoods. Neighborhoods with lower local employment access, be it due to distance or other factors (measured by η_j) are persistently likely to also face large barriers to out-migration (see Table 2).

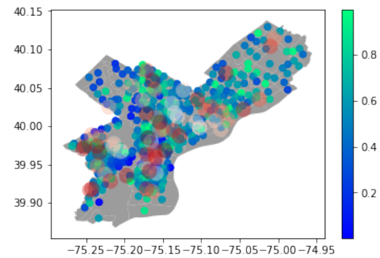
Variable	Estimator	p-value
$\{f_j \| \eta_j < Q_1^{\eta_j}\}$ - $\{f_j \| \eta_j > Q_3^{\eta_j}\}$	1.845	0.07

Table 2: Normalized and dimensionless test statistic comparing inertia in lowest and highest quartiles of local employment mobility based on estimated posterior.

What about the community effect? As the title of this dissertation suggests, there is no degree of certainty that the unexplained pull factor represented by inertia is a financial issue. However, a measure of community is beyond the scope of this paper. Sufficient for this analysis is to demonstrate the correlation of this inertia measure with the presence of key institutions which form communities. Take, for example, charter schools for elementary-aged children. Such schools can develop strong connections between parents and, for young families, these connections provide an amenity that very few economic measures can accurately capture. I explore this experiment here. In figure (16), I map the locations of Philadelphia charter schools.



(a) Locations of elementary charter schools around the City of Philadelphia. Colors correspond to rating by parents combined with test scores. Courtesy of GreatSchools API.



(b) Charter schools with a buffer of 0.01 miles, overlaid with inertia estimates by census tract.

Figure 16: Posterior analysis of how inertia can capture community effects by focusing on charter elementary schools.

The mapping makes it rather clear that charter schools generally tend to be in areas where pull factors are being underestimated. People want to stay in areas with charter schools, and this is not something that is originally in our posterior analysis. The hypothesis that proximity to charter elementary schools have a significant effect on local inertia has a t-statistic of 2.983 resulting in a significance at the 98th percentile. There are some concerns with this analysis at first pass, not least of which that proximity to charter schools tend to correlate with areas of high income. I plot the distribution of income by group in (17). In fact, in Philadelphia, it appears that a majority of the areas with charter schools tend to have an average median income. Few low-income tracts have charter schools. Focusing just on the middle-income charter schools and creating an income cutoff at $80,000, however, increases the significance of the t-test rather than diluting it. This seems to indicate that the income effect alone is not responsible for generating the pattern these rudimentary comparative statistics suggests.

6 Discussion

The figure (18) maps out the distribution of f around Philadelphia. Notice that mobility traps tend to coincide heavily with areas with limited income. Key spatial behaviors are confirmed; for example, inertia tends to grow as distance from key commercial centers like center city Philadelphia increases. Neighborhoods

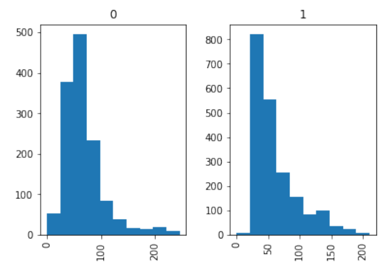


Figure 17: Histogram of tract mean income in thousands of dollars by presence of a charter school (0 = no school in the tract).

with the highest inertia are those one may expect; Kingsessing, West Kensington, Germantown, and Juniata Park, Eastwick, and Roxborough. Manayunk poses an interesting case to one versed in Philadelphia consumption patterns, as it is present as a potential mobility trap despite potentially high income shoppers spending time in the area. It is middling on access to employment. Perhaps this is explained by a large amount of economic vitality being rooted in a few small businesses along the main streets of Manayunk, surrounded by an older mill town. Most of these neighborhoods have been characterized as blighted or low-income. Center city and cultural centers like South Philadelphia tend to have higher fixed costs combined with higher employment access, which could be explained by community resettlement costs.

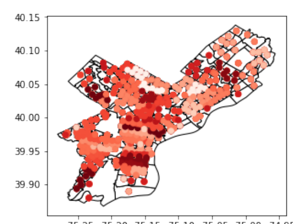


Figure 18: Map demonstrating values of the intertia to mobility variable across Philadelphia census tracts.

I argue that generally speaking, the inertia measure seems to include both noneconomic forces and economic forces. The noneconomic piece I interpret as being in line with non-measurable social structures and community effects. Such inertia relates to a loss of local competitive advantages accrued through cultural or economic agglomeration benefits if a resident were to leave - rather than pure payments alone. These are still relevant considerations in the latent variable, and can be estimated by the residual values from the regression in Table (1).

Next, I consider a direct comparison of neighborhoods which have undergone revitalization before the sample period and those which are marked as suffering from neighborhood blight at any point in the period. I focus on neighborhoods which have undergone revitalization to emphasize that the neighborhoods all have some history of being assessed as poverty traps, but certain neighborhoods are selected as treated areas by Philadelphia planning revitalization at a given time. Identification is threatened due to racial politics of blight having changed over the course of the last 30 years, so I only compare neighborhoods labeled after 1995. Looking at estimated fixed costs at the end of the analysis window in 2016 in Table 3, the differences are clear.

Revitalized Neighborhoods	0.623755
Neighborhoods Certified Blighted	0.546623
T-test Statistic	5.011

Table 3: Comparison of blighted neighborhood fixed costs before and after revitalization by Philadelphia Planning.

Areas with the highest inertia included 2nd & Oxford, the naval base (which, since 2016, has been rapidly redeveloping), and East Overbrook. These areas have radically high inertia. The comparison fixed effect cannot be ascribed directly to a change in mobility costs or a change in community quality, however, which poses a threat to identification of either effect. In particular, it is possible that both are responsible and the revitalized neighborhoods are actually very much suffering from lower rates of local community formation rather than an easier fluidity of leaving the neighborhood. However, considering some of these revitalized neighborhoods - Fishtown and Center City, for example - the community effects are likely what have inflated the inertia estimator. A significant consideration is the potential missing variables problem. If there is a force that is not present in this model, outside of prices, aggregated amenities, fixed costs, housing components, and employment access, then it may diminish the purported influence of fixed costs or alter measurement tactics of local employment access.

The implications of this study are in many ways a measurement claim, where the construction process for inertia was an important exercise. Using a Bayesian latent variable construction to measure what econometricians usually assign to fixed effects allows parsing where the conventional econometric measurement paradigm fails for the low-income individual. It also provides a clear estimation avenue for community, indirect, and noneconomic effect size on fundamental economic processes. These are the key results of the paper aside from the implication about the role of fixed effects in creating mobility patterns in low-income neighborhoods.

7 Conclusion

The rate of movement into neighborhoods has been studied in critical applications surrounding gentrification, neighborhood decline, and place-based policy. Understanding measures of stability and who leaves the neighborhood, however, is a critical method for understanding who bears the cost of changing neighborhood landscapes. This work attempts to develop replicable methodologies for studying neighborhood out-migration as a means to studying geographic mobility. In doing so, we find that using the geographic concepts of push and pull factors to describe amenities with disamenity and amenity value, respectively, is useful for developing a cohesive model with full specification of elasticities.

The model suggests that the variables used in a vast majority of papers on neighborhood choice do not necessarily explain the entirety of the calculus of leaving a neighborhood. In fact, the "usual culprits" *overpredict* out-migration and therefore either underpredict the relative "pull factors" or overpredict "push factors" of the neighborhood. In some neighborhoods around the test case of Philadelphia, the severity of this prediction error is upwards of 60%. I refer to this prediction error as residential inertia.

Two sources of potentially un-measured variation in the model come from variable credit availability across neighborhoods and varying community strength across the same. I test both of these after extracting a cleaned measure of residential inertia. I find that both explanations hold water, both by testing a regression of financial variables on neighborhood inertia (which led to an R^2 of 35%) and by looking at income-controlled comparisons between neighborhoods with and without anchoring institutions like charter elementary schools.

If nothing else, this work becomes a useful exercise in using a simple hierarchical Bayesian data generating process to develop measures for unexplained variation. This method has its risks, but the use of histogram-based migration data curtails some concerns about overfitting individual-level behavior. The model presents some useful conclusions about developing an economic version of a push-pull framework while also presenting a unique way to see how neoliberal models of neighborhood choice still have work to do to prove an empiricist's case.

Further, this paper's results opens the door to a more dynamic model which considers potential duration dependence effects on individuals who do not leave when their locally-derived utility falls below a competitive rate. These individuals may then face deterioration of property values in the long term which could lead to the larger poverty trap cycle. Such a dynamic setting was left for future work due to limitations of the publicly available data used.

References

S. Athey, D. Blei, R. Donnelly, and F. Ruiz. Counterfactual inference for consumer choice across many product categories. Working Paper, 2017a.

S. Athey, D. Blei, R. Donnelly, F. Ruiz, and T. Schmidt. Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. Working Paper, January 2017b.

I. Barakova, R. W. Bostic, P. S. Calem, and S. M. Wachter. Does credit quality matter for homeownership? *Journal of housing Economics*, 12(4):318–336, 2003.

N. Baum-Snow and D. A. Hartley. Accounting for central neighborhood change, 1980-2010. Working Paper, December 2017.

P. Bayer, R. McMillan, A. Murphy, and C. Timmins. A dynamic model of demand for houses and neighborhoods. *Econometrica*, 84(3):893–942, 2016.

P. Bayer, F. Ferreira, and S. L. Ross. What drives racial and ethnic differences in high-cost mortgages? the role of high-risk lenders. *The Review of Financial Studies*, 31(1):175–205, 2017.

R. Benabou. Equity and efficiency in human capital investment: the local connection. *The Review of Economic Studies*, 63(2):237–264, 1996.

R. Benabou. Unequal societies: Income distribution and the social contract. *American Economic Review*, 90(1):96–129, 2000.

M. Betancourt. A conceptual introduction to hamiltonian monte carlo. *arXiv preprint arXiv:1701.02434*, 2017.

D. G. Blanchflower, P. B. Levine, and D. J. Zimmerman. Discrimination in the small-business credit market. *Review of Economics and Statistics*, 85(4):930–943, 2003.

H. Bleakley and J. Lin. Portage and path dependence. *The quarterly journal of economics*, 127(2):587–644, 2012.

W. A. Brock and S. N. Durlauf. Discrete choice with social interactions. *The Review of Economic Studies*, 68(2):235–260, 2001.

R. Chetty, N. Hendren, P. Kline, E. Saez, and N. Turner. Is the united states still a land of opportunity? recent trends in intergenerational mobility. *American Economic Review*, 104(5):141–47, 2014.

R. Chetty, D. Grusky, M. Hell, N. Hendren, R. Manduca, and J. Narang. The fading american dream: Trends in absolute income mobility since 1940. *Science*, 356(6336):398–406, 2017.

V. Couture and J. Handbury. Urban revival in america, 2000 to 2010. Technical report, National Bureau of Economic Research, 2017.

M. A. Davis, J. Gregory, D. A. Hartley, and K. Tan. Neighborhood choices, neighborhood effects and housing vouchers. 2017.

S. N. Durlauf. Neighborhood feedbacks, endogenous stratification, and income inequality. 1993.

S. N. Durlauf. A theory of persistent income inequality. *Journal of Economic growth*, 1(1):75–93, 1996.

R. Ewing, S. Hamidi, J. B. Grace, and Y. D. Wei. Does urban sprawl hold down upward mobility? *Landscape and Urban Planning*, 148:80–88, 2016.

P. Fader and J. Lattin. Accounting for heterogeneity and nonstationarity in a cross-sectional model of consumer purchase behavior. *Marketing Science*, 12 (3):304–317, 1993.

M. Halme and M. Kallio. Likelihood estimation of consumer preferences in choice-based conjoint analysis. *European Journal of Operational Research*, 239(2):556–564, December 2014.

J. J. Heckman and B. E. Honore. The empirical content of the roy model. *Econometrica: Journal of the Econometric Society*, pages 1121–1149, 1990.

Q. Huang, D. Parker, T. Filatova, and S. Sun. A review of urban residential choice models using agent-based modeling. *Environment and Planning B: Urban Analytics and City Science*, 41(4):661–689, 2014. ISSN 2399-8083. doi: 10.1068/b120043p.

J. Hwang and J. Lin. What have we learned about the causes of recent gentrification? *Cityscape*, 18(3):9–26, 2016.

Y. M. Ioannides and L. Datcher Loury. Job information networks, neighborhood effects, and inequality. *Journal of economic literature*, 42(4):1056–1093, 2004.

P. Jehiel and L. Lamy. A mechanism design approach to the tiebout hypothesis. *Journal of Political Economy*, 126(2), 2018 2018.

S. Lee and J. Lin. Natural amenities, neighbourhood dynamics, and persistence in the spatial distribution of income. *The Review of Economic Studies*, 85(1): 663–694, 2017.

J. Lin et al. Understanding gentrification's causes. *Economic Insights*, 2(3): 9–17, 2017.

J. J. Louviere. Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information. *Journal of Choice Modeling*, 1(1):128–164, 2008.

T. Morikawa, M. Ben-Akiva, and D. McFadden. Discrete choice models incorporating revealed preferences and psychometric data. *Advances in Econometrics*, 16:29–55, 2002.

A. I. Mueller, J. Spinnewijn, and G. Topa. Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence and bias. Technical report, National Bureau of Economic Research, 2018.

J. Rothwell and D. Massey. Geographic effects on intergenerational income mobility. *Economic Geography*, 91(1):83–106, 2015.

E. Royall. *Towards an epidemiology of gentrification: modeling urban change as a probabilistic process using k-means clustering and Markov models*. PhD thesis, Massachusetts Institute of Technology, 2016.

J. C. Stein. Prices and trading volume in the housing market: A model with down-payment effects. *The Quarterly Journal of Economics*, 110(2):379–406, 1995.

A. Tranfaglia. Shrinking networks: A spatial analysis of bank branch closures. Working Paper 19-12, Federal Reserve Bank of Philadelphia, March 2018.

B. Van der Klaauw and J. C. Van Ours. From welfare to work: does the neighborhood matter? *Journal of Public Economics*, 87(5-6):957–985, 2003.

M. Walls, N. Magliocca, and V. McConnell. Modeling coastal land and housing markets: Understanding the competing influences of amenities and storm risks. *Ocean and Coastal Management*, 157:95–110, 2018.

A Events Controls

Event	Significant at 95%
Passage of 2015 Lower Court Ruling preventing City from forcing developers to maintain vacant buildings	No
Recession 2009-2011	Yes

B Coefficients in Churn Model, No Fixed Costs

Variable	α_{jt} Estimator	β_{jt} Estimator	Significance at 95%
Distance, Groceries	0.534		Y
Distance, Gasoline	-0.044		Y
Distance, Bank Branch	0.362		Y
Distance, Nightlife and Entertainment	0.029		Y
Train accessibility	1.228		Y
Bus accessibility	0.805		Y
Ratings for Local Groceries	0.077		N
Ratings for Local Entertainment	0.039		Y
Local school ratings (Google)		0.199	Y
Local crime rates	0.138		Y
Local Income Variability		0.185	Y
Racial evenness estimator	0.090		N
Income exposure index		0.068	Y
η_{jt}	0.225		Y
Housing price-market value residual		0.740	Y
Housing hedonic value	0.165		N
Event fixed effect	0.050	0.023	N
Vacant lot evenness		0.552	Y
Commercial plate evenness	0.007	0.035	N
Vehicle ownership	0.100		Y

Table 4: Results of beta-geometric regression without a fixed cost estimator. Neither was significant, though the estimated effect size in the pull estimator was stronger (at a similar significance level of about $p < 0.25$).