Location Choice and the Marginal Utility of Consumption

Md Moshi Ul Alam
Clark University

Morris A. Davis
Rutgers University

Jesse Gregory
University of Wisconsin

March 20, 2025

Abstract

We address a fundamental identification challenge in optimal place-based redistribution policy design: characterizing how marginal utility of consumption (MUC) varies across locations. We implement a nationally representative novel survey eliciting location-specific MUC by asking respondents to evaluate versions of themselves across different potential locations. Our findings reveal that the average MUC at the city level is weakly but positively correlated with average local income. Furthermore, using exogenously varied location-specific income, our survey elicits direct measures of location-specific MUC. The patterns in the data reject the common specification in location-choice models that consumption utility is additively separable from idiosyncratic location attachment shocks. We develop a structural model with imperfect substitutability between consumption and the idiosyncratic shocks to explain the empirical patterns. We find that optimal policy is substantially more nuanced than a simple redistribution from high-wage to low-wage locations.

1 Introduction

Economic models of spatial equilibrium provide a framework for studying optimal taxation and redistribution across locations. When these models incorporate externalities and a motive for across-group redistribution (often implemented through Pareto weights), the planner's problem determines taxes and transfers to achieve two main objectives: ensuring that agents internalize the externalities of their location choices and redistributing across groups to balance efficiency and equity. A third, often less emphasized, consideration for policy design is the variation in the marginal utility of consumption (MUC) across locations within a given type–such as individuals with the same skill level. Just as redistribution across different types is a key motive, this within-type variation in MUC also creates an incentive for redistribution.

This paper focuses on this third force, within-type variation across places in the marginal utility of consumption. We demonstrate that MUC cannot be identified using location choice data alone, which poses a challenge for standard models of spatial policy. To address this limitation, we design and implement a nationally representative survey to directly elicit location-specific MUCs. Using a hypothetical choice experiment, we ask respondents to consider multiple versions of themselves living in different locations and identify which version would benefit the most from a permanent and unexpected income increase. This approach allows us to estimate MUC at the individual and city level and assess its relationship with income across geographic areas.

To investigate how the marginal utility of consumption (MUC) varies across locations, we designed and implemented a nationally representative survey that directly elicits location-specific

MUCs. The core question asks respondents to imagine multiple versions of themselves, each living in a different potential future location, and to identify which version would benefit the most from a permanent and unexpected increase in income. This approach allows us to measure MUC within individuals, rather than relying on comparisons across individuals, which would confound MUC with unobserved differences in preferences or characteristics. We only aggregate across individuals to assess which cities, on average, have higher or lower MUCs.

To further understand how MUC varies with local income levels, we present respondents with a sequence of hypothetical scenarios that manipulate income while holding other factors constant. In each scenario, the respondent considers the same set of potential future locations but with different assigned income levels. This allows us to observe how the location in which they report the highest MUC shifts as income in different locations changes. By systematically varying income while keeping amenities, social networks, and other location attributes fixed, we isolate the effect of income differences on location-specific MUC.

We summarize the data from the survey responses in several ways to examine how MUC varies across individuals and locations. In the baseline scenario, where households compare locations assuming their income in each place is equal to what they expect to earn there, we find that within individuals, MUC tends to be higher in higher-wage locations. This also implies that, on average, MUC is higher in higher-wage places. However, we find that differences in average income across cities explain little of the overall variation in MUC across locations. This weak overall correlation could arise either because income directly affects utility in ways that vary across locations or because high-wage locations differ in other ways that influence MUC.

To disentangle these explanations, we analyze responses to the scenario questions, in which respondents are explicitly told that all factors other than income remain unchanged. This ensures that the variation in income across locations is exogenous, allowing us to isolate the effect of income on MUC. We find that an exogenous increase in income within a location raises the average MUC reported for that location. This pattern suggests two distinct forces that do not arise in separable models of location choice: a direct effect and a selection effect. The direct effect follows from standard consumption theory—marginal utility of consumption decreases as income rises, holding other factors constant. The selection effect, however, operates through changes in the composition of individuals living in a location. When a location's income rises, individuals with weaker attachment to that location become more willing to move there. Whether these new arrivals have higher or lower MUC than incumbents depends on the extent to which location attachment is substitutable or complementary with consumption. If attachment and consumption are substitutes, new arrivals will tend to have higher MUC than incumbents, raising the location's average MUC. If they are complements, new arrivals will have lower MUC, reducing the average.

Using the empirical results we reject the standard model of location choice for optimal place based policy that assumes consumption utility is additively separable from the idiosyncratic location specific shocks. We show that this is because the additively separable model by construction does not allow for any selection effect conditional on location choice. Next, we develop a structural

model that allows for imperfect substitutability between consumption and the idiosyncratic shocks to explain the empirical patterns. The imperfect substitutability allows marginal entrants to a given location to have different marginal utilities of consumption than inframarginal entrants. The exogenous variation that we induce in the survey allows us to identify the parameters of the model. We estimate the model using Indirect Inference (Gourieroux et al., 1993) and find that the model is able to replicate the empirical patterns in the data by being able to match all the target moments within numerical tolerance. Using the estimated model we can simulate counterfactual policies and optimal policy.

Our results demonstrate that the linearly separable model, which implies transfers from high to low-wage cities, produces a welfare loss equivalent to 0.06 percentage points lower consumption. While we caution against overinterpreting our initial attempt at building the above complexities in the model, our findings suggest that optimal policy should consider not only wages, but also person and location-specific characteristics and their substitutability with wages. This more nuanced approach increases welfare by the equivalent of 0.065 percentage points higher consumption.

Our paper is organized as follows. In Section 2, we characterize optimal policy and show that it is not identified in standard models of location choice using location choice data. Next we outline our survey design and highlight the key features of our survey that allow us to elicit location-specific MUC essential for evaluating place-based policies and designing optimal policy. Then we describe various patterns in our data related to location choice elasticity as well as the location-specific MUC. We conduct extensive validation checks against publicly available data sources to demonstrate that respondents' subjective assessments meaningfully correspond to objective location characteristics documented in the literature. After showing that our estimates on the elasticity of location choice are consistent with documented estimates, we describe the empirical patterns in the marginal utility of consumption across locations and how they relate to the average income in those locations. Next, we present a structural model of location choice that allows us to explain the empirical patterns in the data followed by its estimation and discussion of the estimates. Finally, we conclude with simulations of optimal policy using the estimated model and discuss the implications of our findings for optimal policy.

2 Theory and Lack of Identification

2.1 Characterizing Optimal Policy

We document the lack of identification of optimal transfers in commonly used models of location choice. Before we discuss identification, we characterize optimal policy. Consider a simple setup with only 2 locations[1] where locations differ in their income per household, denoted w_1 and w_2. Assume residents of location 2 pay a tax t to finance a subsidy b paid to residents of location 1. Each household draws idiosyncratic preferences for locations we label as ε_1 and ε_2 and chooses the location that provides the highest utility. Utility is derived from consumption bundled with each

[1] It is straightforward to extend what follows to an N-location framework.

individual's draws of ε_1 and ε_2. Expected utility in this simple model is

$$V = E_{\varepsilon_1,\varepsilon_2} \max\Big(u\big(w_1+b,\varepsilon_1\big), u\big(w_2-t(b),\varepsilon_2\big)\Big) \quad \text{with} \quad t(b) = \frac{L_1}{1-L_1}b$$

where L_1 is the measure of households living in location 1 and the second expression is the government balanced budget condition assuming there are a measure 1 of households in the economy.

The optimal subsidy is the value of b at which $\frac{dV}{db} = 0$. To characterize this optimal subsidy, begin with the expression,

$$\frac{dV}{db} = \frac{\partial V}{\partial b} + \frac{dt}{db}\frac{\partial V}{\partial t} + \Big(\frac{\partial L_1}{\partial b} + \frac{dt}{db}\frac{\partial L_1}{\partial t}\Big)\underbrace{\frac{\partial V}{\partial L_1}}_{=0}$$

The third term is equal to zero because individuals who switch locations in response to a marginal policy change receive the same utility in both locations.[2]

The partial effects of changing b and t depend only on the populations of the two locations and the average marginal utility of consumption of residents in each location, denoted μ_1 and μ_2:

$$\begin{aligned}
\frac{\partial V}{\partial b} &= L_1 \overbrace{E\left[\frac{\partial u\big(w_1+b,\varepsilon_1\big)}{\partial w_1}\middle| u\big(w_1+b,\varepsilon_1\big) > u\big(w_2-t,\varepsilon_2\big)\right]}^{\equiv\mu_1} \\
\frac{\partial V}{\partial t} &= -(1-L_1)\underbrace{E\left[\frac{\partial u\big(w_2-t,\varepsilon_1\big)}{\partial w_2}\middle| u\big(w_1+b,\varepsilon_1\big) < u\big(w_2-t,\varepsilon_2\big)\right]}_{\equiv\mu_2}
\end{aligned}$$

The government balanced budget condition implies that

$$\frac{dt}{db} = \frac{L_1}{1-L_1}\left(1 + \frac{\frac{dL_1}{db}\frac{b}{L_1}}{1-L_1}\right)$$

The first term is the mechanical change in the tax that is necessary to finance the change in transfers in the absence of any behavioral responses. The second term captures the fact that an additional tax increase is necessary to offset the population response to the change in taxes/transfers. We can now write

$$\frac{dV}{db} = L_1\mu_1 - \frac{L_1}{1-L_1}\left(1 + \frac{\frac{dL_1}{db}\frac{b}{L_1}}{1-L_1}\right)(1-L_1)\mu_2 = L_1\left[\mu_1 - \mu_2\left(1 + \frac{\frac{dL_1}{db}\frac{b}{L_1}}{1-L_1}\right)\right]$$

In the case when the current transfer is $b = 0$, the planner wishes to transfer consumption to the location with the higher marginal utility of consumption. At the optimal transfer b (satisfying

[2] dt/db is not a partial derivative both because people move as a result of the policy and because of the need to balance the budget.

$dV/db = 0$), we have

$$\frac{\mu_1 - \mu_2}{\mu_2} = \frac{\frac{dL_1}{db}\frac{b}{L_1}}{1 - L_1} \tag{1}$$

This result is exactly analogous to the Baily-Chetty formula of Chetty (2006) characterizing the optimal generosity of unemployment insurance. The left-hand side is the insurance benefit of moving 1 dollar (in total) from location 2 to location 1. The right-hand side describes the marginal cost of raising one dollar in total from location 2, which captures the fact that as benefits rise, the population of location 1 also increases which causes an excess burden of transferring one additional dollar.

2.2 Proving Lack of Identification

Equation (1) makes clear that optimal policy requires knowledge of the average marginal utility of consumption in each location and the elasticity of location choice with respect to taxes and transfers. Economists can directly infer the elasticity of location choice with respect to taxes and transfers (Rossi-Hansberg et al., 2022). The marginal utility of consumption is not directly observable.

To make progress, researchers specify (i) a utility function for consumption and place-based attachment and (ii) a distribution for place-based attachment and then use data on incomes and optimal choices to infer how the marginal utility of consumption varies across locations. So far, two such specifications span the literature. In the first, the utility for household i choosing to live in location n is $a_n c_n e_{in}$, where a_n are freely available amenities common to everyone in location n, c_n is consumption per household of all households in location n, and e_{in} is household i's attachment to location n which is drawn from the Fréchet distribution. In the second, the utility function is $\ln a_n + \ln c_n + e_{in}$ where e_{in} is drawn from the Type 1 Extreme Value Distribution. In both of these specifications, the left-hand side of equation (1) can be shown to be equal to

$$\frac{1/c_1 - 1/c_2}{1/c_2}$$

Both specifications predict optimal policy transfers income from high-income locations to low-income locations. Any differences that arise are attributable to different modeling choices affecting the right-hand side of equation (1).

Unfortunately, neither the distribution of the attachment draws nor the exact way in which the attachment draws should enter utility is identified by available data. Researchers therefore make arbitrary and untestable assumptions on the distribution of the attachment draws and how the draws enter utility and changing these assumptions can change predicted optimal policy.

Consider a two-location model and a specification for utility in each location of

Location:	Utility	Marginal Utility
1	$\ln(c_1 + e_{i1})$	$1/(c_1 + e_{i1})$
2	$\ln(c_2)$	$1/c_2$

Assume that $c_2 > c_1$ and that e_{i1} is drawn from a distribution with support strictly larger than $-c_1$. In this specification, household i chooses to live in location 1 whenever $e_{i1} > c_2 - c_1$ and therefore for all households choosing to live in location 1, $c_1 + e_{i1} > c_2$. Thus, the average realized marginal utility of consumption in location 1, $1/(c_1 + e_{i1})$ is always less than that of location 2, $1/c_2$. In this framework, a policy-maker optimally transfers income from low-income location 1 to high-income location 2, exactly the opposite result of existing papers.

Conditional on a specification for how attachment draws enter utility, these models suffer from a similar but related identification problem. Consider a utility function of consumption in location n and location-specific preferences of household i for location n as $u(c_n, e_{in})$. Denote $n^*(i)$ as the optimal location choice among $n = 1, \ldots, N$ locations. That is, given values of $e_{i1}, e_{i2}, \ldots, e_{iN}$,

$$n^*(i) = \text{argmax}\Big[u(c_2, e_{i1}), u(c_2, e_{i2}), \ldots, u(c_N, e_{iN})\Big]$$

Now for each household, introduce a random variable D_i that multiplicatively affects utility. For any given household the value of D_i does not change with that household's location and therefore D_i does not change optimal location choices.

$$n^*(i) = \text{argmax}\Big[D_i u(c_2, e_{i1}), D_i u(c_2, e_{i2}), \ldots, D_i u(c_N, e_{iN})\Big]$$

D_i is not estimable from location-choice data or incomes, since it does not change any decisions.

If D_i is correlated with e_{ni}, then optimal policy can be anything. To see this, consider a two location example where utility is $D_i c_n e_{in}$ and e_{in} is drawn from the Fréchet. Now consider three cases for D_i that show different correlations of D_i with e_{i1} and e_{i2}. In all three cases, D_i does not affect location choice:

		Utility		Marginal Utility	
Case	D_i	1	2	1	2
1	$D_i = 1$	$c_1 e_{i1}$	$c_2 e_{i2}$	e_{i1}	e_{i2}
2	$D_i = 1/e_{i1}$	c_1	$c_2 (e_{i2}/e_{i1})$	1	e_{i2}/e_{i1}
3	$D_i = 1/e_{i2}$	$c_1 (e_{i1}/e_{i2})$	c_2	e_{i1}/e_{i2}	1

In all three cases, household i chooses location 1 whenever $c_1 e_{i1} > c_2 e_{i2}$ and location choice data does not identify D_i.

Optimal transfers depend on the specification of D_i. To see this, consider the case $c_1 = c_2$, implying location 1 is chosen whenever $e_{i1}/e_{i2} > 1$ and location 2 is chosen whenever $e_{i2}/e_{i1} > 1$.

Case 1. $D_i = 1$. Optimal policy involves no transfers, see Rossi-Hansberg et al. (2022).

Case 2. $D_i = 1/e_{i1}$. The marginal utility of consumption of households in location 1 is always 1, but the marginal utility of consumption of households choosing location 2 is $e_{i2}/e_{i1} > 1$. Optimal policy requires consumption to be transferred from location 1 to location 2.

Case 3. $D_i = 1/e_{i2}$. The marginal utility of consumption of households in location 2 is always 1, but the marginal utility of consumption of households choosing location 1 is $e_{i1}/e_{i2} > 1$. Optimal policy requires consumption to be transferred from location 2 to location 1.

We now wish to explain the economic content of D_i. D_i is something fundamental to the household that affects the marginal utility of consumption. The presence of D_i does not need to directly affect location choice; for D_i to influence predictions on optimal policy, it simply needs to be correlated with location choice. Setting $D_i = 1$ for all households is an arbitrary assumption that may not be supported by data and may lead to completely incorrect predictions for optimal policy. The following simple example illustrates this point. Suppose people can either be in (unobservably) good health or poor health, and assume the marginal utility of consumption is relatively high for people in poor health. Now consider a two location model, where location 1 has an extreme climate and location 2 has a mild climate. People in bad health may have relatively low preferences for living in a location with an extreme climate. This scenario corresponds to our Case 2, where D_i is negatively correlated with e_{i1}. Case 2 calls for optimal transfers from location 1 (extreme climate) to location 2 (mild climate). Ultimately, this occurs because people in poor health and a high marginal utility of consumption are more likely to live in location 2, a place with a mild climate.

To recap, we have shown optimal place-based policies satisfy equation (1). We have also shown that researchers must take a stand on the distribution of location-specific preferences, on how these preferences enter the utility function, and on all other variables that may affect the marginal utility of consumption and may be correlated with the draws of the location-specific preferences. In our first paper, we show that existing data on location choices and wages can not identify any of these variables. We demonstrate this using theory and numerical examples, first with simple models such as we have discussed here, but then in more complicated models with arbitrarily many locations and externalities in production.

To make progress on optimal policy that takes into account differences in the marginal utility of consumption, we propose to measure the marginal utility of consumption under various scenarios using a hypothetical-choice surveys, described next, that measures variation in the marginal utility of consumption across different locations.

3 Data

As demonstrated in Section 2, location choice data alone cannot identify location-specific MUC, as multiple utility specifications can generate identical choice probabilities while implying divergent redistribution policies. This fundamental non-identification problem motivates our methodological approach. We address this challenge through a novel survey instrument specifically designed to elicit both subjective location choice probabilities and the MUC across locations conditional on

location choice. The key innovation in our survey is the direct measurement of location-specific MUC through structured hypothetical choice experiments, enabling us to estimate the average MUC across locations and thereby infer optimal place-based redistributive policy. We focus on the primary features of the survey while leaving the details on design and implementation to Appendix A.

3.1 Eliciting choice probabilities and the marginal utility of consumption

The key part of our survey designed to elicit choice probabilities and MUC is divided into their two broad respective parts. Before this, each respondent is asked to report their hometown where they grew up and three other locations that they think they are most likely to live in the next 5 years. Next, respondents are asked to report their location-specific expected annual income, presence of family and friends, their ranking of perceived expenses and ease of pursuing their hobbies and habits, and the probability of choosing each location in the next 5 years. We call this to be the baseline scenario.

Then we present respondents with a series of scenarios where we exogenously shock their reported location-specific expected income and presence of family and friends, and ask them to report their choice probabilities across the four locations. This completes the first part of the survey. In the second part of the survey, we ask the respondent to imagine different counterfactual worlds where they have chosen to live in each of the four locations. Then we ask them that suppose they receive an unanticipated permanent income shock, in which location would that permanent jump in income be the most valuable to them. Once they choose a location, we ask them to report the compensating differential of a permanent income shock that would make them just as happy as they would be in the chosen location. We repeat this exercise over the same set of exogenously varied scenarios as before and ask them to report their compensating differentials.

Thus, for each respondent, we have a panel data with 24 linearly independent observations on choice probabilities and 24 linearly independent observations on compensating differentials in response to exogenous shocks to their location-specific incomes and presence of family and friends, in addition to their baseline responses.

3.2 Key highlights of our survey design

One major difference that our survey design has compared to the literature using hypothetical choice surveys for identification is that our identification exploits variation within each location between scenarios for each individual and not within scenarios which imposes stronger assumptions on unobservable location-specific heterogeneity. Specifically, we ask individuals to assume that given a location, any unobserved factor remains the same across scenarios. Thus, we can allow for unobserved differences across the locations which are held constant across scenarios. This is weaker than assuming that everything else between the locations within each scenario is held fixed, thereby instructing the respondents to make choices based on the differences in the observed characteristics of the locations. Additionally, by allowing for unobserved differences across locations, we reduce

the cognitive load on respondents, and make the identification strategy more robust to variations in unobserved factors. For completeness, we are able to do this because the choice sets involve MSAs which are well-defined unlike other surveys that use hypothetical choice experiments to elicit preferences for objects in choice-sets which can have varied definitions and thus can be more difficult to define and compare, for example hypothetical jobs.[3]

We also differ in the way we induce experimental variation in our survey. Instead of providing quasi-randomly drawn incomes, we draw shocks following the distribution in income documented in Blundell et al. (2008) and use those shocks to exogenously vary their expected location-specific income across different scenarios. To exogenously vary the expected presence of family and friends we quasi-randomly draw zeros and ones for each of the four locations.[4] The advantage of this approach is that the exogenously shocked incomes are specific to the individual-location pair and are thus more realistic than the quasi-randomly drawn incomes for each of the four locations.

3.3 Other data sources for validation tests

We supplement our survey data with location-specific characteristics from three other sources primarily for validation checks. The first source is the 2022 American Community Survey (ACS), providing demographic and economic indicators including median house prices, wage indices, and housing cost measures. The amenity indices developed by Diamond (2016), which quantify location-specific amenity indices including school quality, retail access, crime rates, congestion measures, environmental quality, and non-wage job attributes, are the second source. The third source is the Zillow Home Value Index (ZHVI), which provides monthly data on median home prices for each MSA.

Using these additional sources of data we conduct multiple validation tests in Section 4.3 to check the internal consistency of our survey responses. We first check whether the expected location-specific income reported by respondents is positively correlated with a skill adjusted location-specific wage index that we construct from the ACS. We also examine the correspondence between subjective rankings of location "expensiveness" and median house prices, and subjective rankings of location "fun-ness" and the amenity indices.

4 Patterns in our Survey Data

Our survey design elicits information about location preferences and the marginal utility of consumption across different locations. We collected responses from 2,193 individuals with non-missing data on key questions, combining data from our second pilot and final data collection.

[3]For example Wiswall and Zafar (2018) use hypothetical choice experiments to identify preferences over hypothetical job attributes, and Alam et al. (2021) embed information experiments within hypothetical choice surveys to identify beliefs on gender gaps in manager quality in hypothetical jobs.

[4]Note that the variations we induce are quasi-random and not completely random. This is done to avoid draws in scenarios which can lead to choices being strictly dominant in all attributes thereby rendering variations in those scenarios to be ineffective following Wiswall and Zafar (2018) and Alam et al. (2021).

4.1 Comparison of our sample to the ACS

We collected demographic information, including age, race, gender, income, and education level. The following figure shows the distribution of respondents by age and income level compared to the American Community Survey (ACS) data.

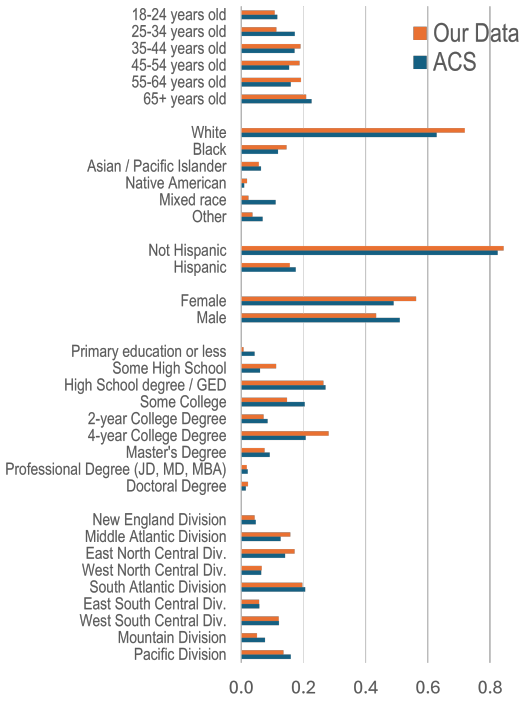


Figure 1: Demographics

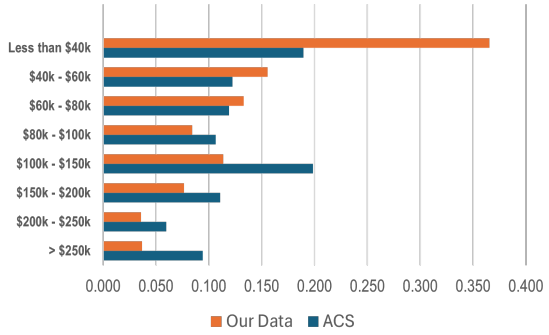


Figure 2: Income

The figure shows that the marginal distributions of demographic variables in our sample are reasonably representative of the U.S. population. However, the joint distributions of these characteristics show some differences compared to the ACS. To address these imbalances in the joint distribution, we implement inverse probability weighting to achieve two-way demographic balance, though results herein use unweighted data. With respect to household income, our sample shows an overrepresentation of lower-income households relative to the ACS distribution.

4.2 Spatial distribution of respondents and their future potential locations

The geographic distribution of respondents covers a wide range of metropolitan areas across the United States as shown in Figure 3. Respondents listed a diverse set of potential future metropolitan areas in Figure 4 with darker shaded MSAs indicating locations that were chosen more than those shaded lighter. The spatial distribution of these destinations reflects both the geographic dispersion of our sample shown in the previous figure and their intended migration patterns. Major coastal metropolitan areas feature prominently, though respondents also identify a substantial number of interior MSAs.

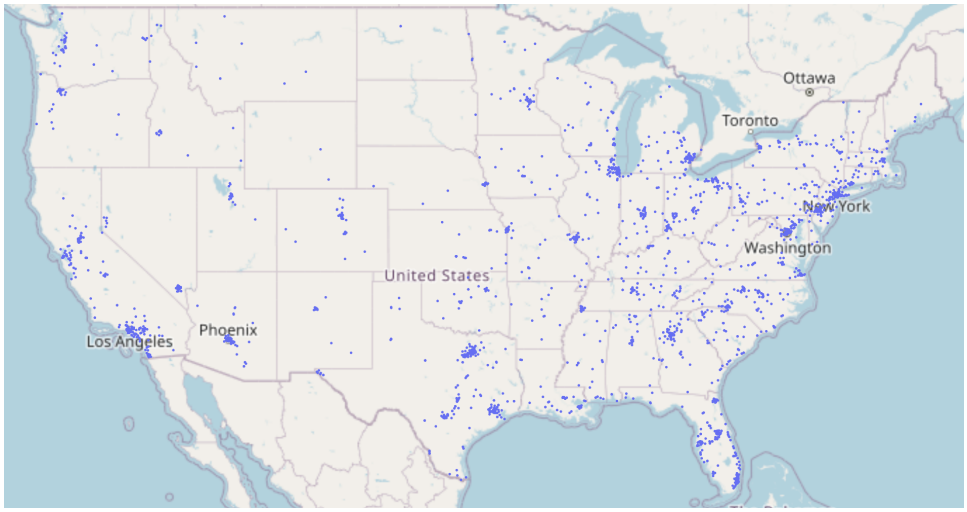


Figure 3: Location of Respondents

4.3 Validation of survey responses

We conduct several validation exercises to confirm that respondents provided substantively meaningful information.

First, we examine the relationship between reported location-specific expected income and metro-level wage index. Specifically, using respondents' baseline responses on their location-specific expected income, we estimate

$$ln(ExpInc_{ij}) = \alpha_0 + \alpha_1 CensusWageIndex_j + \mu_i + e_{ij}$$

where μ_i is a respondent fixed effect and the Census wage index is the skill-adjusted hourly wage index for the location j that we estimate from the American Community Survey (ACS).[5]

[5]To construct a location-specific wage index adjusted for skill, we first regress log hourly wages on demographic characteristics (age, education, race, sex, and marital status) and metropolitan area fixed effects using the 2022 ACS 5-year sample, restricting to full-time workers (working more than 35 hours per week). The resulting metropolitan area fixed effects represent skill-adjusted average hourly wages that account for demographic composition differences across locations, providing a measure of location-specific wage premiums.

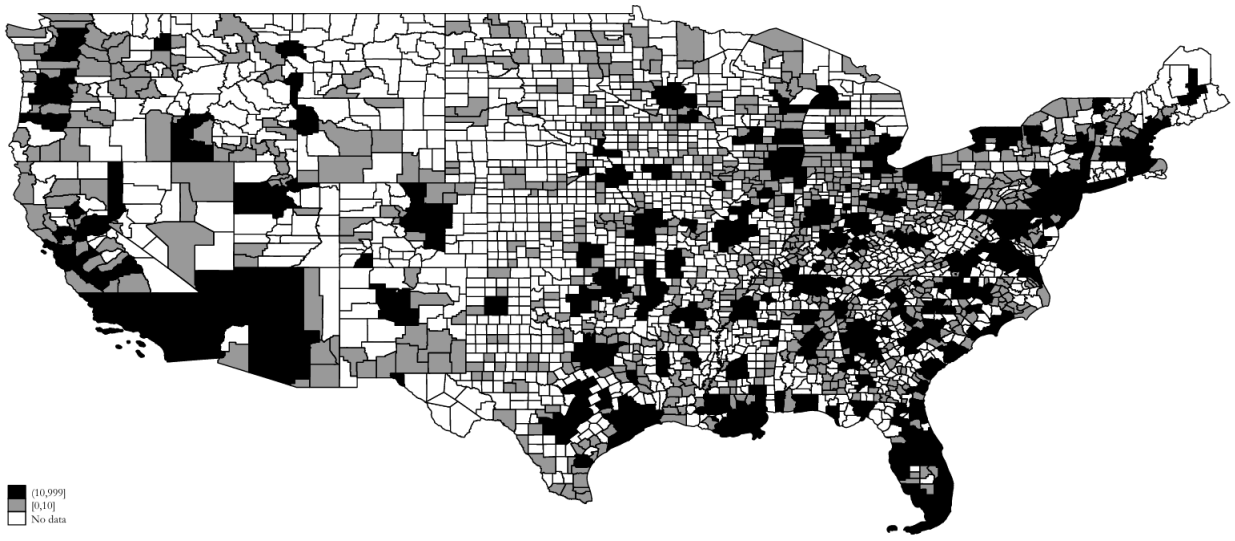


Figure 4: Potential Future Locations

Dependent Variable:	
Log(expected household income)	
Variable	**Coefficient**
Census Wage index (ACS)	0.421
	(0.017)
Constant	4.330
	(0.001)
Observations	64,503
R^2	0.842
Individual FEs	Yes

Robust standard errors in parentheses.

Our estimates reveal a statistically significant positive relationship between expected income and skill-adjusted metro hourly-wages. Indeed, respondents anticipate higher incomes in metropolitan areas with higher wage indices. This indicates that respondents anticipate systematic income differentials across metropolitan areas in patterns that align with observed wage variation, suggesting that income expectations align with regional labor market conditions.

Next, we show that reported probabilities of any move in the next five years strongly align with established life-cycle patterns, with moving probabilities declining monotonically with age shown in the binned scatterplot of Figure 5. Young adults (18-24) report approximately 60% probability of moving within five years, declining systematically to approximately 20% for respondents over 65. This pattern is consistent with existing literature on migration patterns, confirming that our survey captures expected life-cycle dynamics in location choice.

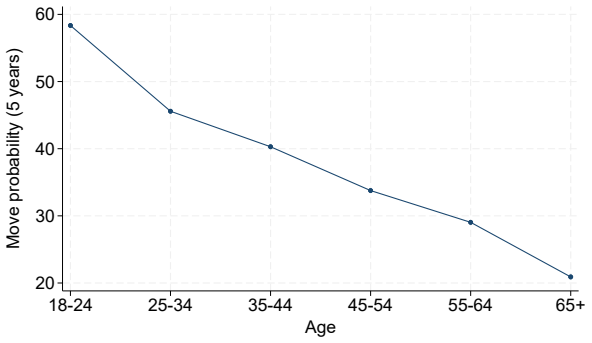


Figure 5: Reported chance of any move in the next 5 years falls with age

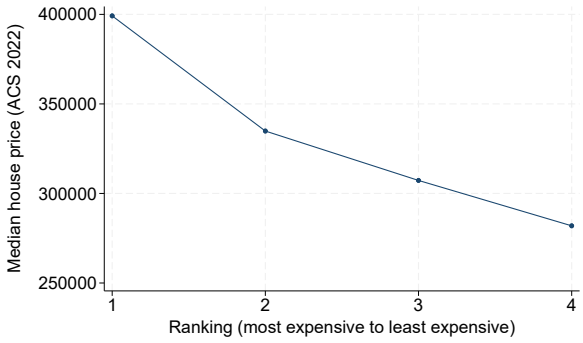


Figure 6: Reported rankings align with metro median house prices

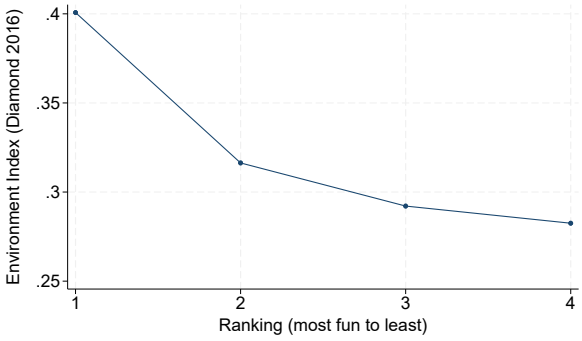


Figure 7: Reported rankings align with park spending / air quality

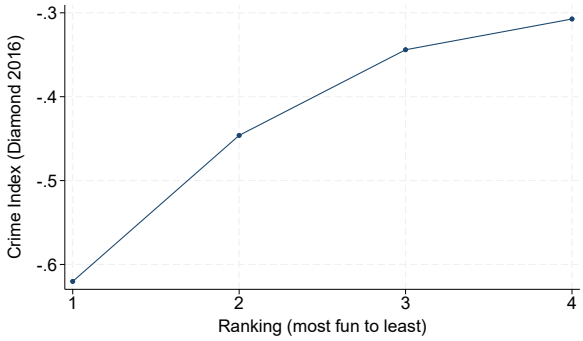


Figure 8: Reported rankings align negatively with metro crime rates

Respondents' rankings of metropolitan area characteristics also show strong correlation with objective measures. Binned scatterplots in Figure 6 reveal that when respondents rank candidate locations by "expensiveness," these rankings show a strong positive correlation with median house prices from ACS data, with median house prices monotonically decreasing from the most to least expensive ranked locations.[6] Similarly, binned scatterplots of rankings of metropolitan "fun-ness" in Figure 7 demonstrate positive association with objective measures of amenities, such as environmental quality indices that incorporate per-capita park spending and air quality measures obtained from Diamond (2016).[7] These same "fun-ness" rankings correlate negatively with metropolitan crime indices obtained from Diamond (2016) and shown in Figure 8, indicating that respondents perceive locations with lower crime rates as more enjoyable.

[6]We ask "Given the sorts of goods and services you like to buy and given the kind of housing unit you wish to live in, which locations do you think would be most expensive for you?" and then ask the respondent to rank the likely future locations from 1 (most expensive) to 4 (least).

[7]We ask "Thinking about your hobbies and things you like to do for entertainment, where do you think you will have the most fun?" and then ask the respondent to rank the likely future locations from 1 (most fun) to 4 (least).

4.4 Elasticities of location choice

The experimental variation we induce over multiple scenarios in our survey allows us to identify the elasticity of location choice with respect to income. Specifically, our instructions in the survey guide respondents to hold location specific unobservable constant across scenarios, while we exogenously vary the expected income and the presence of family and friends in each of the four locations across eight scenarios. Thus, conditional on an individual-location fixed effect and an individual-scenario fixed effect, variation in expected income and the presence of family and friends across scenarios allows us to identify the elasticity of location choice with respect to income and the presence of family and friends.

We note that choice probabilities are typically responded in multiples of five and thus have measurement error associated with them; we use a least absolute deviation estimator (Blass et al., 2010).[8] Specifically, we consider the following reduced form specification:

$$\log(p_{ijk}) = \nu_1 \log(w_{ijk}) \; + \; \nu_2 f_{ijk} \; + \; \phi_{ik} \; + \; \phi_{ij} \; + \; u_{ijk} \tag{2}$$

where p_{ijk} is the location choice probability of household i in location j in scenario k, log income is $\log(w_{ijk})$), f_{ijk} is an indicator variable for the presence of family and friends, ϕ_{ik} are person-scenario fixed effects and ϕ_{ij} are person-location fixed effects.

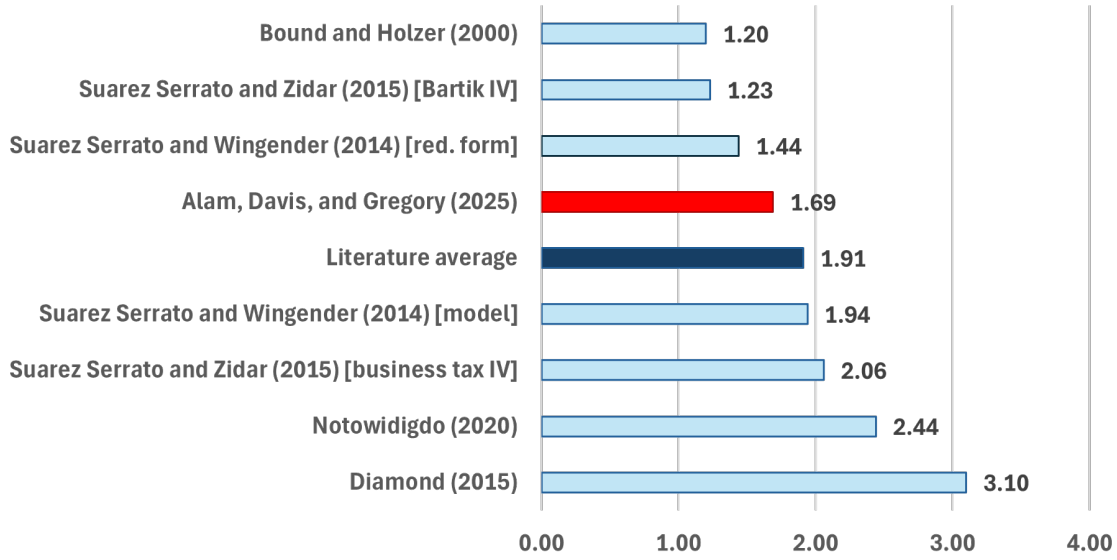
Taking into account measurement error in reported choice probabilities, and using experimental variation of our survey allows identification through the assumption of $\mathbb{M}[u_{ijk} \mid w_{ijk}, f_{ijk}, \phi_{ik}, \phi_{ij}] = 0$. However, since the conditional median independence assumption conditions on fixed effects as well, we estimate the above equation using a two-step quantile regression approach following Canay (2011). In the first step we estimate the fixed effects ϕ_{ik} and ϕ_{ij} using a simple fixed effect estimator for equation (2), and residualize the estimates $\widehat{\phi}_{ik}$ and $\widehat{\phi}_{ij}$ from the log choice probabilities $log(p_{ijk})$. In the second step we estimate the quantile regression of the residualized log choice probabilities on the experimentally induced variation in income and the presence of family and friends. The key assumption in this two-step estimator is that the fixed effects are simply shifters of all quantiles by the same amount. Then as the number of scenarios approaches infinity, the quantile regression of the residualized log choice probabilities on the experimentally varied income and the presence of family and friends yields consistent estimates of the elasticities of location choice with respect to income and the presence of family and friends. We block bootstrap the standard errors at the individual level.

Our analysis yields an estimated location choice elasticity with respect to income (ν_1) of 1.69 (SE = 0.004), indicating substantial responsiveness of location decisions to income differences.[9] Our estimate is squarely in line with existing estimates from the literature, which typically range in between 1.2 and 3.1 depending on context and methodology as shown below.

[8] The estimator has an attractive property that it is not sensitive to small perturbations of extreme probabilities of 0 and 1.

[9] Our estimate of the location elasticity with respect to the presence of family and friends is $\nu_2 = 0.20$ with a standard error of 0.007, quantifying the non-monetary value of social connections in location choice.

Figure 9: Comparison of our estimates to the literature estimates synthesized in Fajgelbaum et al. (2019)



To estimate the empirical distribution of the elasticities across respondents, we run the above estimation routine separately for each individual i. Indeed, as scenarios for each individual approach infinity, the individual location choice elasticities are identified. However, since estimation uses 8 scenarios which gives us 24 linearly independent observations for each individual, the standard errors will be very high. Thus, to estimate the distribution of location choice elasticity we implement a parametric empirical Bayes (also known as a shrinkage estimator) approach which shrinks the individual estimates towards the mean of the distribution based on estimated signal-to-noise ratio. We explain the details of shrinkage procedure with the two-step estimator described above in Appendix B.[10] The graph below plots the shrinked distribution of location choice elasticities after implementing the empirical Bayes approach. The distribution is centered around 1.73, with a standard deviation of 0.55 which yields a standard error for the mean equal to 0.011.

5 Empirical Patterns in the Marginal Utility of Consumption

In order to explain the patterns in the marginal utility of consumption across locations, we first show how the survey elicits data that can be mapped to the MUC within the context of a generic choice model. Hence, before we describe the empirical patterns in the location-specific MUC, we first provide a micro-foundation for construction of variables which can be used as measures of location-specific MUC.

[10] Also see Walters (2024) for an excellent review on parameteric and non-parametric empirical Bayes methods and application in labor economics.

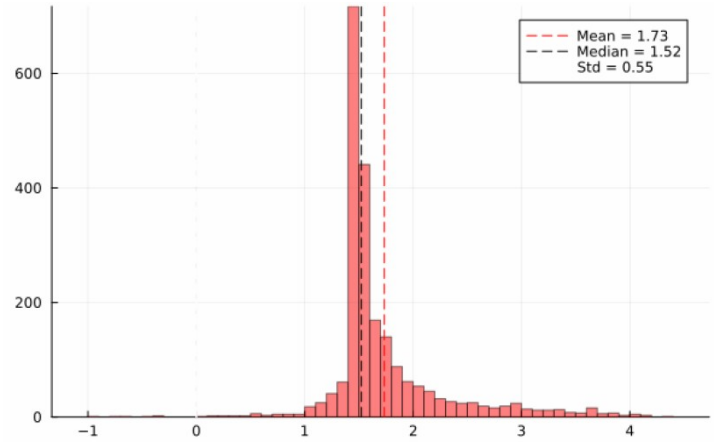


Figure 10: Distribution of location choice elasticity estimates

5.1 Variable construction to utilize MUC

For the purposes of this section, abstract from the presence of family and friends. Denote the utility for household i from location j as

$$U_{ij}\left(C_{ij}, X_j, \epsilon_{ij}\right)$$

where C_{ij} is consumption, X_j are amenities of the location common to all people, and ϵ_{ij} are person-specific amenities.[11] The marginal utility of consumption is

$$MUC_{ij} \quad = \quad \frac{dU_{ij}\left(C_{ij}, X_j, \epsilon_{ij}\right)}{dC_{ij}}$$

When asked about location where MUC is highest, we envision respondents comparing expected MUCs conditional on choice:

$$E\left[MUC_{ij} \mid j \text{ chosen}\right] \quad = \quad E_{\epsilon_{ij}}\left[\frac{dU_{ij}\left(C_{ij}, X_j, \epsilon_{ij}\right)}{dC_{ij}} \mid u_{ij} = \max\left\{u_{ij'}\right\}\right]$$

The average of the marginal utility of consumption depends on the (a) the level of consumption C_{ij}, (b) the complementary of consumption with X_j and with ϵ_{ij}, and (c) the mix of ϵ_{ij} for which j is the chosen location. The correlation between the average of the marginal utility of consumption and the level of consumption in any given location can therefore be positive or negative.

When respondents answer this question:

> *In which location would you benefit the most from an unexpected and permanent increase of $5,000 in household income, had you chosen to live there?*

we assume the respondent is making comparisons of the expected marginal utility of consumption,

[11] The presence of X_j allows the marginal utility of consumption to vary across locations that offer the same consumption.

conditional on choice. In other words, we assume the answer (j^*) satisfies this condition:

$$E\left[\frac{dU_{ij^*}\left(C_{ij^*}, X_{j^*}, \epsilon_{ij^*}\right)}{dC_{ij^*}} \mid j^* = \text{argmax}\right] \quad > \quad E\left[\frac{dU_{ij}\left(C_{ij}, X_j, \epsilon_{ij}\right)}{dC_{ij}} \mid j = \text{argmax}\right] \quad \forall j \neq j^*$$

Then, when respondents answer this question,

> *Now ... indicate the size of the jump in household income that they would need living in those locations 5 years from now such that their change in happiness after receiving this jump in household income is the same as the change in happiness from the $5,000 jump in household income ... in [location of previous answer].*

we assume the response for each location $j \neq j^*$, which we denote as b_j, satisfies the following condition:

$$E\left[\frac{dU_{ij}\left(C_{ij}, X_j, \epsilon_{ij}\right)}{dC_{ij}} \mid j = \text{argmax}\right] b_{ij} \quad = \quad E\left[\frac{dU_{ij^*}\left(C_{ij^*}, X_{j^*}, \epsilon_{ij^*}\right)}{dC_{ij^*}} \mid j^* = \text{argmax}\right] \$5,000$$

For each household i, for all $j = 1, \ldots, 4$, and for any given scenario $k = 0, \ldots, 8$ (where $k = 0$ is the baseline) this enables us to compute for each j

$$\frac{1}{b_{ijk}} \quad = \quad \zeta_{ik}\, E\left[\frac{dU_{ij}\left(C_{ij}, X_j, \epsilon_{ij}\right)}{dC_{ij}} \mid j = \text{argmax}\right] \tag{3}$$

where ζ_{ik} is a person- and scenario-specific constant. In other words, for each person and for each scenario, we can identify the marginal utility of consumption in three of the likely locations relative to the fourth. This implies that if we define $muc_{ijk} = 1/b_{ijk}$, then if we regress $log\left(muc_{ijk}\right)$ on person-by-scenario fixed effects, we can uncover how the marginal utility of consumption varies by location and directly estimate how income and other factors correlate with the variation.

5.2 Cross-Sectional Analysis of Marginal Utility of Consumption

To estimate the variation in marginal utility of consumption (MUC) across locations, we implement two complementary approaches using our cross-sectional baseline data. The first approach non-parametrically estimates metropolitan-specific MUC values in an AKM framework (Abowd, Kramarz, and Margolis 1999), while the second examines systematic relationships between MUC and observable location characteristics using data from Diamond (2016) and the American Community Survey (ACS).

For our non-parametric estimation, we specify:

$$\ln(muc_{ij}) = \theta_i + \theta_{metro} + e_{ij} \tag{4}$$

where $\ln(muc_{ij})$ represents the log of the inverse required income boost in location j (relative to the highest-MUC location) for respondent i, θ_i captures individual fixed effects, and θ_{metro} represents metropolitan area fixed effects that constitute our parameters of interest. This approach, analogous

to the AKM (Abowd, Kramarz, and Margolis 1999), allows us to obtain metropolitan-specific MUC estimates while controlling for respondent heterogeneity. Using estimates of θ_{metro}, we regress them on skill adjusted average hourly wage indices from the ACS described above to obtain a measure of the marginal utility of consumption in each metropolitan area.

Our second approach examines the relationship between MUC and observable metropolitan characteristics:

$$\ln(muc_{ij}) = Z_j'\gamma + \theta_i + e_{ij} \tag{5}$$

where Z_j' represents a vector of metropolitan characteristics. This specification enables us to analyze systematic patterns in how MUC varies with location attributes.

5.2.1 Cross-Sectional Results

Our non-parametric estimates reveal substantial heterogeneity in average MUC across metropolitan areas. We show the geographic variation in the estimated metropolitan fixed effects θ_{metro} in Figure 11 below.

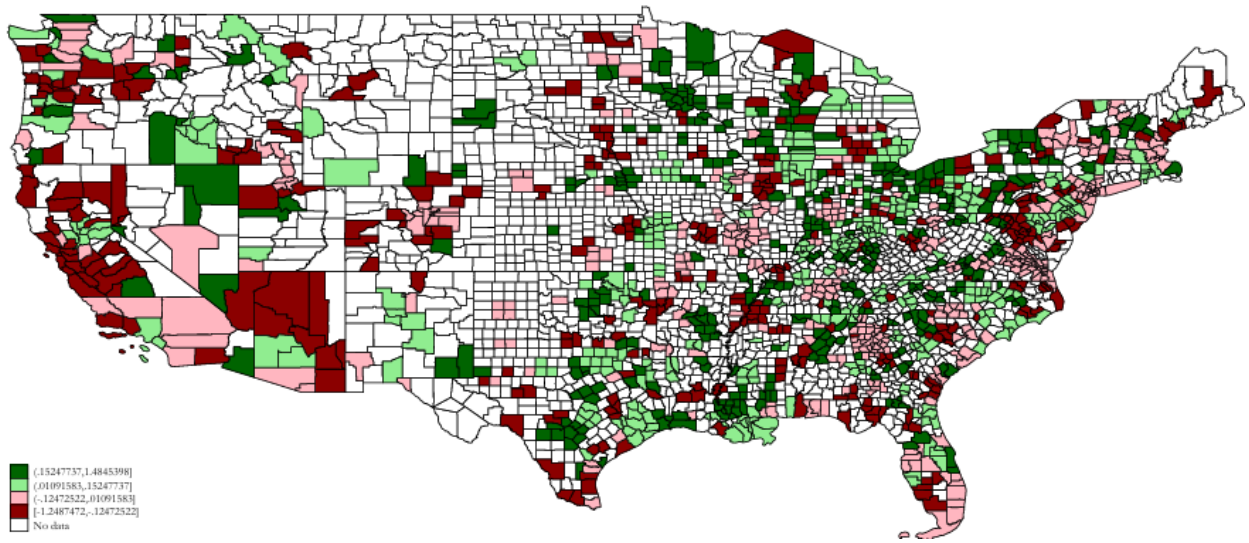


Figure 11: Estimated metropolitan fixed effects θ_{metro} from AKM model

When examining the relationship between MUC and metropolitan characteristics, we find multiple significant associations reported in Table 1. Our full-sample specification (column 1) yields an estimated elasticity of MUC with respect to expected household income of 0.197 (SE = 0.0058), indicating that marginal utility of consumption is significantly higher in locations with higher expected income. Concurrently, we find a negative elasticity with respect to median house prices of -0.091 (SE = 0.0061), suggesting that after controlling for income, locations with higher housing costs exhibit lower marginal utility of consumption. When restricting to metropolitan areas covered by Diamond's (2016) amenity indices (columns 2-4 of Table 1), the income elasticity increases to 0.244 (SE = 0.0085), while the housing price elasticity remaining similar at -0.105 (SE = 0.0100).

In our most comprehensive specification, we find that school quality (coefficient = 0.0160, SE = 0.0051) and non-wage job amenities (coefficient = 0.0110, SE = 0.0039) are positively associated with MUC, while retail amenities (coefficient = -0.0113, SE = 0.0046) and crime (coefficient = -0.0010, SE = 0.0042) show negative associations. Traffic indexes (coefficient = -0.0012, SE = 0.0041) and air quality (coefficient = 0.0005, SE = 0.0039) show no significant relationship with MUC though the signs are consistent with our expectations. The final column of Table 1 shows the results from the AKM model with individual fixed effects, wherein we report the estimate from regressing estimated metro fixed effects on skill adjusted wage indices constructed from the ACS, which yields similar estimates.

5.3 Panel Analysis of Marginal Utility of Consumption

5.3.1 Experimental Design and Estimation

To identify causal relationships between income and marginal utility of consumption, we leverage our experimental design where respondents evaluated multiple hypothetical scenarios with varying income levels and presence of family/friends across locations. This panel structure enables us to include multiple layers of fixed effects that control for unobserved heterogeneity.

For each respondent i in each scenario k, we construct $\log(MUC_{ijk}) \equiv \log(1/b_{ijk})$ for each location j, where b_{ijk} represents the income boost required in location j to match the utility gain from a \$5,000 income boost in the highest-MUC location. Our estimating equation is:

$$\log(MUC_{ijk}) = \beta_1 \log w_{ijk} + \beta_2 f_{ijk} + \theta_{ik} + \theta_{ij} + e_{ijk} \tag{6}$$

where θ_{ik} and θ_{ij} represent individual-by-scenario and individual-by-location fixed effects, respectively. Our experimental variation allows us to use this specification to isolate the causal effect of income changes (and changes in presence of family and friends) on marginal utility of consumption conditional on individual-location and individual-scenario fixed effects.

5.3.2 Panel Results

Our panel analysis yields an estimated elasticity of MUC with respect to expected income (β_1) of 0.45 (SE = 0.01), substantially larger than our cross-sectional estimates. This difference likely reflects the experimental design's ability to isolate causal effects from confounding factors. The presence of family and friends (β_2) increases MUC by approximately 4% (SE = 0.00), indicating a complementarity between social connections and the marginal utility of consumption.

To assess heterogeneity in these relationships, we estimate respondent-specific elasticities using:

$$\log(MUC_{ijk}) = \beta_{i1} \log w_{ijk} + \beta_{i2} f_{ijk} + \theta_{ik} + \theta_{ij} + e_{ijk} \tag{7}$$

We implement an empirical Bayes shrinkage procedure to adjust for sampling variance, shrinking individual-specific estimates toward the sample mean based on the estimated signal-to-noise ratio.

Table 1: Cross-Sectional Analysis of Marginal Utility of Consumption

	Individual level regressions			Metro level regressions
	Our data	Diamond (2016) data		AKM
	Log(MUC)	Log(MUC)	Log(MUC)	$\hat{\theta}_{metro}$
Log(Expected HH income)	0.0762***	0.107**	0.102**	
	(0.0229)	(0.0383)	(0.0383)	
ACS log wage index				0.133
				(0.198)
Log(Median house price)	-0.0852***	-0.0941**	-0.169***	-0.165**
	(0.0193)	(0.0352)	(0.0443)	(0.0573)
School Index			0.0294	
			(0.0164)	
Retail Index			0.00379	
			(0.0149)	
Crime Index			-0.0402**	
			(0.0143)	
Road Index			-0.00168	
			(0.0128)	
Environment/Parks Index			0.0146	
			(0.0161)	
Jobs Index (non-wage)			0.0146	
			(0.0129)	
Constant	-8.462***	-8.465***	-7.509***	2.077**
	(0.247)	(0.453)	(0.565)	(0.731)
Observations	8772	4856	4856	93
Individual FE	Yes	Yes	Yes	N/A
R^2	0.355	0.476	0.479	0.136

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The distribution of these estimates reveals substantial heterogeneity in how respondents' marginal utility of consumption responds to income changes across locations, as shown in Figure 12. The mean elasticity is 0.43 with a standard deviation of 0.16 yielding a standard error of 0.003, indicating considerable variation in the strength of this relationship.

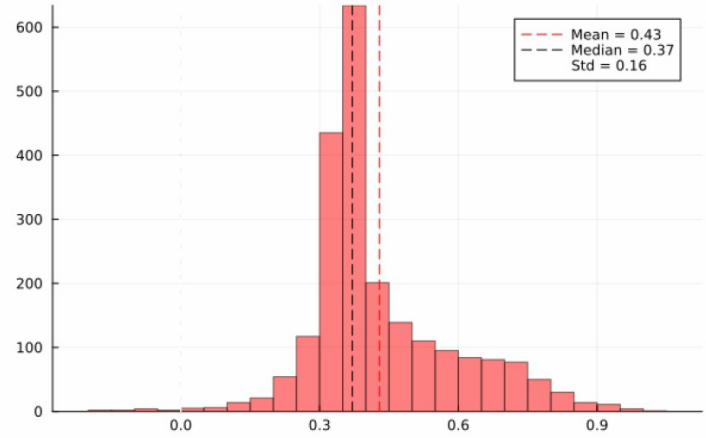


Figure 12: Distribution of Estimates of MUC Elasticity

The panel estimates inform the theoretical relationship between income and MUC conditional on location choice described above. As shown earlier, the expected marginal utility of consumption in location j given that j is chosen can be expressed as:

$$\frac{d}{dw_j}E[MUC_j|j \text{ chosen}] = \underbrace{\int \frac{\partial MUC_j(w_j, e_i)}{\partial w_j} f(e_i|j \text{ chosen}) de_i}_{\text{Direct effect}} + \underbrace{\int MUC_j(w_j, e_i) \frac{\partial f(e_i|j \text{ chosen})}{\partial w_j} de_i}_{\text{Selection effect}} \tag{8}$$

The positive overall elasticity of 0.45 indicates that selection effects dominate direct effects, with the latter typically negative in standard models of risk aversion. This pattern suggests that marginal entrants to higher-income locations have systematically higher MUC than inframarginal residents. In standard linearly separable models the selection effect by construction is zero thereby not allowing any heterogeneity between inframarginal and marginal individuals. These findings have significant implications for optimal redistribution policies across locations, as they demonstrate that the standard approach of simply redistributing from high-income to low-income locations may not align with welfare maximization when accounting for heterogeneity in marginal utility of consumption.

6 Model-Based Results

Next, to investigate what underlying preferences and distributions of location attachment are consistent with our main empirical results, and to study their implications for optimal redistributive

policy, we develop and estimate a simple structural model. In this preliminary draft, we study a stylized model of a homogeneous population (and thus a choice problem represented by a single random utility specification) that we estimate by targeting key average elasticities and sample moments.

6.1 Rejection of separable model

Several of our main empirical results findings are difficult or impossible to rationalize with discrete location choice model that specifies consumption utility as additively separable from location-specific shocks or attachment draws, such as:

$$\log(w_{jk}) + \epsilon_{ij} \tag{9}$$

where w_{jk} represents income in city j for skill group k, and ϵ_{ijk} are location specific attachment draws that captures individual-specific location preferences. This model places strong restrictions on how MUC varies across locations (on average) and across individuals within a location. First, MUC is a (decreasing) function of the consumption level (wage). As a result, high wage cities must have lower average MUCs, and within a city individuals with the same wage will have the same MUC regardless of the strength of their attachment to the city. These model predictions do not change when the utility specification is augmented to allow non-wage factors to affect location choice by including additive linear terms in observable location characteristics or including an additive composite location amentity parameter (a location fixed effect).

Our main empirical findings reject these restrictions implied by the additively separable model. Specifically, the standard additively separable model cannot rationalize the findings that:

- The average MUC of each location conditional on the location being chosen is *positively* correlated with the location's wage level, but variation in wage levels across locations explains a small fraction of the overall variation in average MUC across locations.
- Within individual, the expected MUC in a location conditional on that location being chosen is *increasing* the location's wage.

The selection effect ($\int MUC_j(w_j, e_i)\frac{\partial f(e_i|j \text{ chosen})}{\partial w_j} de_i$) in the separable model is zero by construction. We introduce an alternative utility function that allows for substitutability between income and person-specific amenities, which generates these patterns in equilibrium allowing for selection effects.

6.2 Non-separable Model and Estimation Strategy

To capture the interactions between income and location-specific amenities, we specify the following utility function:

$$\frac{\left(w_{ijk}^{\rho} + (\lambda\ a_{ij}\ \varepsilon_{ij})^{\rho}\right)^{\frac{1-\sigma}{\rho}}}{1-\sigma} \tag{10}$$

where:

- ρ governs the elasticity of substitution between income and person-specific amenities.
- σ represents the coefficient of relative risk aversion.
- λ scales the effect of person-specific amenities.
- a_{ij} are location-specific factors that vary across individuals but remain fixed for each individual in a given location.

This specification introduces the flexibility needed to allow differences in MUC across locations, even among individuals with similar skills.

We estimate ρ, σ, and λ by targeting key moments in the data. For each respondent, we standardize income across locations by dividing their reported income in a given location by their highest reported income and multiplying by 100. This allows income comparisons across individuals.

6.3 Model

6.3.1 Locations

In the model, the agent i considers four locations $\mathbf{j = 0, 1, 2, 3}$.

- $\mathbf{j = 0}$: denotes the agent's hometown.
- $\mathbf{j = 1, 2, 3}$: denotes the three most likely alternative locations ranked from highest to lowest wage.

6.3.2 Factors affecting location-specific utility:

Location-specific utility is given by $u_{ijk} = u(w_{ijk}, a_{ij}, \varepsilon_{ijk})$. The factors entering utility are:

- **Consumption level:** equal to the wage $\mathbf{w_{ijk}}$ available in each location (k indexes scenarios across which the realized wages vary)
- **Permanent location characteristics/attachments:** denoted by $\mathbf{a_{ij}}$, this value describes the agents valuation of amenities in location j and aspects of the agent's attachment to location j that are known at the time of the survey.
- **"Shocks":** denoted by $\varepsilon_{\mathbf{ij}}$ described by changes to life circumstances and location attachments revealed between survey and the time the agent makes its location choice.

6.3.3 Timing and information:

There are two stages of the model:

- **Survey stage:** 5 years prior to location choice.

- The agent has partial information $\Omega_0 = \{w_{ijk}, a_{ij}\}_{j=0}^3$ about the factors determining location-specific utility.
- Other payoff-relevant information Ω_1 is not yet known, but the agent knows the distribution Ω_1 will be drawn from.

- **Location choice stage:**
  - The agent has full information $\Omega = \{\Omega_0, \Omega_1\}$ about the factors determining location-specific payoffs.
  - $\Omega_1 = \{\varepsilon_{ij}\}_{j=0}^3$

- **Implications for linking model with survey responses:** Because households have not yet observed the ε component of location attachment at the time they answer the survey, location choices and MUCs conditional on choice are random variables at the time of the survey. We interpret respondents' reported probabilities of location choice and expected marginal utility conditional on choice as expectations taken over possible location draws with Ω_0 known.

6.3.4 Location choice:

Let $j_k^* = \arg\max_j u(w_{ijk}, a_{ij}, \varepsilon_{ijk})$ be the location chosen in the second stage.

6.3.5 Marginal Utility of Consumption:

For a given realization of the location attachment draws, the agent's marginal utility of consumption in their chosen location is:

$$MUC_{ij^*k} = w_{ij_k^*k}^{\rho-1} \left(w_{ij_k^*k}^{\rho} + \left(\lambda \; a_{ij_k^*} \; \varepsilon_{ij_k^*} \right)^{\rho} \right)^{\frac{1-\sigma}{\rho}-1} \tag{11}$$

At the time of the survey, the expected marginal utility of consumption in a specific location j conditional on j being chosen is,

$$E\left[MUC_{ijk} \mid j \text{ chosen}\right] = E_{\varepsilon}\left[MUC_{ij_k^*} | j_k^* = j\right] \tag{12}$$

$$E\left[MUC_{ij} \mid j \text{ chosen}\right] \quad = \quad E_{\epsilon_{ij}}\left[\frac{dU_{ij}\left(C_{ij}, X_j, \epsilon_{ij}\right)}{dC_{ij}} \mid u_{ij} = \max\left\{u_{ij'}\right\}\right]$$

7 Estimation

We estimate the model by Indirect Inference (Gourieroux et al., 1993). The procedure involves choosing a set of targeted moments, denoted by a column vector $\widehat{\gamma}$ containing the values of these moments estimated from the data, and searching for the set of model parameters $\Theta = [a_1, a_2, a_3, \rho, \sigma, \lambda]'$

that result in the same moments estimated using data simulated from the model, denoted by $\tilde{\gamma}(\Theta)$, to be as close as possible to the targets $\widehat{\gamma}$. That is, the estimator of the structural parameters Θ is,

$$\widehat{\Theta} = \arg\min_{\Theta} \; (\tilde{\gamma}(\Theta) - \widehat{\gamma})'(\tilde{\gamma}(\Theta) - \widehat{\gamma}) \tag{13}$$

7.1 Simulation details

For a given candidate vector of model parameters Θ, we compute model analogs of the targeted moments ($\tilde{\gamma}(\Theta)$) using the following steps:

1. Draw R vectors of location draws from a $pareto(1, 0.75)$ distribution.[12]
2. Using the baseline location wages, evaluate the location utility in all four locations for each vector of attachment draws. For each replication $r = 1, ..., R$, record the agent's optimal location choice and MUC.
3. For each location, take averages across replications to find each location's average choice probability and average MUC conditional on the location being chosen.
4. Repeat steps 2 and 3 using the modified wages for each scenario 1-8.
5. Using the simulated choice probabilities and MUCs conditional on choice, compute $\tilde{\gamma}(\Theta)$.

7.2 Targeted moments

For estimation we target the following list of 6 data moments, referred to in the context of indirect inference as auxiliary models:

7.2.1 Key estimated elasticities:

1. **Elasticity of location choice** with respect to income (from equation 2, $\widehat{\nu}_1 = 1.69$).
2. **Elasticity of average location MUC conditional on choice** with respect to income (from equation (), $\widehat{\beta}_1 = 0.45$).

7.2.2 Risk aversion

We discipline the exercise by targeting a coefficient of relative risk aversion 1.0, within the conventional range studied in the literature.

3. **Elasticity of MUC** with respect to income, i.e. the partial effect of income on MUC holding location choices fixed ($\partial \ln MUC_{ij} / \partial \ln w_{ij} = -1.0$).

[12] The choice of 1 is a normalization, since the parameter λ that multiplies the draws plays an identical role to the scale parameter of the distribution. We plan to eventually estimate the shape parameter of the pareto distribution by targeting moments related to the distribution of reported MUCs across locations, but for now we have fixed the value of the shape parameter at 0.75.

7.2.3 Average location incomes and choice probabilities

4, 5, 6. **Location choice probabilities:**

Location	Income Rank	Avg. Income ($,000s)	Choice probability
Non-Hometown	1	97.15	22.11%
	2	85.44	20.39%
	3	75.59	19.25%
Hometown		81.51	38.25%

7.3 Parameter Estimates and Fit to Data

Our estimation procedure is able to match all the target moments within numerical tolerance yielding the following parameter values:

Parameter	Estimate
a_0 (normalization)	1.0
a_1 (highest-income non-hometown)	0.41501
a_2	0.49743
a_3 (lowest-income non-hometown)	0.57606
ρ (elasticity of substitution)	0.54307
σ (coefficient of relative risk aversion)	1.52595
λ (scale of amenities)	11.19160

Table 2: Structural Parameter Estimates

Observe that $\rho = 1-\sigma$ is equivalent to the linearly separable model. The parameters a_j for $j \neq 0$ less than 1 reflect that they are chosen with lower probabilities than location 0 (home-town).

8 Counterfactual Policies and Policy Implications

We investigate the welfare effect of two policies. The first is the welfare effect of the policy implied by the standard linearly separable model, which implicitly assumes that average marginal utility is proportional to $\frac{1}{C}$. This policy implies a tax of $\frac{1}{1+\nu_1}$ which is then redistributed lump-sum resulting in a transfer from high-wage cities to low-wage cities. The results show that the linearly separable model will imply the policy that will transfer from high wage cities to low wage cities. However, this policy will result in a welfare loss equivalent to 0.06 percentage points lower consumption.

The second policy we investigate is the welfare maximizing policy derived from our model's structural parameter estimates subject to two constraints: the tax transfer is uniform within city and balanced budget. While we caution against overinterpreting this first attempt at model building, our findings highlight a key insight: optimal spatial policy is more complex than commonly

			City tax rate	
			Standard model policy	Optimal policy
City	Wage	Prob	tax = $\frac{wage_j - \overline{wage_j}}{1+\nu_1}$	
0	81.51	38.25	-1.040	-3.16
1	97.15	22.11	4.173	0.88
2	85.44	20.39	0.270	2.11
3	75.59	19.25	-3.013	3.96
CE welfare change			-0.06 pct. points	0.065 pct. points

Table 3: Optimal policy and Comparison with other policies

assumed. Our results suggest that optimal transfers should not be based simply on wages alone, nor on choice probabilities alone. Rather, optimal policy must account for how person-and-location-specific characteristics are substitutable with wages to affect the marginal utility of consumption across locations, and how likely are locations to be chosen. Consequently, optimal policy increases welfare equivalent to 0.065 percentage points of higher consumption.

9 Conclusion

Our paper addresses a fundamental identification challenge in optimal spatial policy design: characterizing how the marginal utility of consumption varies across locations. We show that the marginal utility of consumption cannot be identified from location choice data alone, necessitating direct elicitation methods. Through our novel survey approach, including exogenous variations, we obtain direct measures of location-specific marginal utility of consumption.

Our empirical findings reveal that the average marginal utility of consumption has a positive but weak correlation with wage levels across metropolitan areas. This pattern stands in contrast to standard specifications in location choice models that assume consumption utility is additively separable from idiosyncratic location attachment shocks. We develop a structural model that incorporates imperfect substitutability between consumption and location attachment to explain the empirical patterns observed in our data. This allows marginal entrants to a location to have systematically different marginal utilities of consumption than the inframarginal individuals.

The model reconciles three key empirical facts that we document from our data: the elasticity of location choice with respect to expected income, the positive elasticity of marginal utility of consumption with respect to expected income, and conventional levels of risk aversion. Our parameter estimates suggest that the reconciliation requires a utility specification with partial substitutability between consumption and location attachment, along with a distribution of location attachment that has sufficiently thick tails.

These findings have substantial implications for optimal spatial policy. Our counterfactual

simulations demonstrate that the linearly separable model, which implies transfers from high to low-wage cities, produces a welfare loss equivalent to 0.06 percentage points of consumption. In contrast, our simple yet more nuanced model that accounts for person and location-specific characteristics and their substitutability with wages increases welfare by the equivalent of 0.065 percentage points higher consumption. This suggests that optimal policy is substantially more complex than simple redistribution from high-wage to low-wage locations.

References

Alam, M.M.U., Mookerjee, M., Roy, S., 2021. Worker-side Discrimination: Beliefs and Preferences—Evidence from an Information Experiment on Jobseekers. Working Paper.

Blass, A.A., Lach, S., Manski, C.F., 2010. Using elicited choice probabilities to estimate random utility models: Preferences for electricity reliability. International Economic Review 51, 421–440.

Blundell, R., Pistaferri, L., Preston, I., 2008. Consumption inequality and partial insurance. American Economic Review 98, 1887–1921.

Bound, J., Holzer, H.J., 2000. Demand shifts, population adjustments, and labor market outcomes during the 1980s. Journal of labor Economics 18, 20–54.

Chetty, R., 2006. A general formula for the optimal level of social insurance. Journal of Public Economics 90, 1879–1901.

Diamond, R., 2016. The determinants and welfare implications of u.s. workers' diverging location choices by skill: 1980-2000. The American Economic Review 106, 479–524.

Fajgelbaum, P.D., Morales, E., Suárez Serrato, J.C., Zidar, O., 2019. State taxes and spatial misallocation. The Review of Economic Studies 86, 333–376.

Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of applied econometrics 8, S85–S118.

Kuziemko, I., Norton, M., Stantcheva, S., Saez, E., 2015. How elastic are preferences for redistribution: Evidence from randomized survey experiments. American Economic Review 105, 1478–1508.

Notowidigdo, M.J., 2011. The incidence of local labor demand shocks. Technical Report. National Bureau of Economic Research.

Rossi-Hansberg, E., Sarte, P.D., Schwartzman, F., 2022. Cognitive hubs and spatial redistribution. Working Paper.

Serrato, J.C.S., Wingender, P., 2014. Estimating the Incidence of Government Spending. Technical Report. Duke University and the International Monetary Fund.

Serrato, J.C.S., Zidar, O.M., 2015. Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms. Technical Report. Duke University and The University of Chicago Booth School of Business.

Walters, C., 2024. Empirical bayes methods in labor economics, in: Handbook of Labor Economics. Elsevier. volume 5, pp. 183–260.

Wiswall, M., Zafar, B., 2018. Preference for the workplace, investment in human capital, and gender. The Quarterly Journal of Economics 133, 457–507.

A The Survey to Identify the Marginal Utility of Consumption across Locations

A Appendix: Survey design

A.1 Details on design choices

Definition of a location: A location in our survey is defined as a metropolitan statistical area (MSA). We obtain the list of all states and MSAs within each state from the U.S. Census Bureau. We provide a drop-down menu of states followed by a drop-down menu of MSAs within the chosen state. At the beginning of the survey, we define a location as a metropolitan area: *Note: A metropolitan area is a large region comprising one or more counties, which includes a city with a significant population and its surrounding areas that are economically interconnected. It is defined by factors such as population size and the extent of linkage among nearby communities in terms of jobs, commuting, and transportation.*

Additionally, the three locations other than the respondent's hometown are required to be unique and different from the current location of the respondent and the location where they grew up.

Choice of language: We administered the survey in English and in Spanish. The Spanish translation was done by experts at the Rutgers Survey Center. We also conducted a pilot survey and a focus group discussion with a small sample of Spanish-speaking individuals to ensure that the translation was accurate and that the survey was understandable.

Displaying the exogenous location-specific attributes: We color coded the direction of the shocks to highlight the incomes to reduce cognitive load on the respondents. This decision was based on tests of whether the color coding had any effect on the responses in two simultaneous rounds of pilot data collection where we found no difference in the responses between the color-coded and non-color-coded versions of the pilot survey, except for shorter times to completion in the color-coded version. Thus, we decided to use the color-coded version in the final survey to reduce the cognitive load on the respondents. In order to avoid the possibility of any parallel being drawn with political affiliation, we specifically avoided using the colors blue and red. To represent

increment (reduction) in income we used green (yellow). We did not color code the presence or absence of family and friends. We could have used the same color coding for the presence of family and friends, but we decided against it to avoid assuming whether the presence of family and friends increases or decreases utility, which in the case of income is more straightforward. In the same spirit of minimizing cognitive load, we showed respondent the shocked income instead of showing both the percentage change and the new income.

Ordering: Since we are identified within individuals, we are not concerned about the order of the locations. However, we randomized the order of the scenarios to avoid any ordering effects.

Facilitating ease in responses: We made several choices in order to facilitate ease in responding to the survey. Here we some of the most important ones:

- We provided a slider for the respondents to report their expected income and the compensating differential. This also avoids the possibility of different respondents using different units of income.
- In each exogenously varied scenario, respondents were shown their baseline probabilities and compensating differentials to facilitate comparison.
- In order to emphasize that their response in the compensating differential questions is only conditional on that location being their choice in 5 years, each option of location was followed with the phrase "*had I chosen to live there in 5 years*"

A.2 Background

The Survey Research Center at Rutgers University has helped us design and implement the survey. Our survey was administered in the July and November of 2024 to a nationally representative sample of adults. The survey was available in English and Spanish, and participants that completed the survey were given small payments in line with typical compensation given to online survey participants. The survey respondents themselves were anonymous, and no personally identifiable data was collected.

The survey started with a small pilot in April of 2024. Based on responses and feedback from the pilot, we adjusted survey questions and then administered the survey to a larger sample as a second pilot (July). Once we reviewed the results from that pilot and decided not to implement any changes, we then launched the full survey in November.[13]

A.3 Locations and Income

We start the survey by asking basic questions about current and previous location and current and expected future incomes.

[13]The second pilot tested whether responses varied depending on whether certain phrases were highlighted or not. The responses did not seem to depend on highlighting. For the full survey, we used the highlighted version.

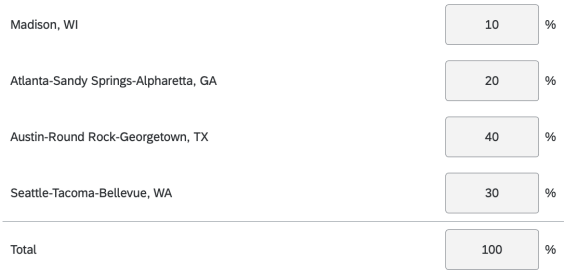
- We first ask the state, metropolitan area or nearest city, and zip code that the respondent considers his or her home town, and then the same data for where the respondent currently lives.
- We next ask the respondent their total pre-tax income over the previous 12 months. If the respondent lives with one or more adults, the respondent is asked to only report income from his or her job.
- Then we ask the respondent to report his or her expected individual annual income 5 years from now.
- Then we ask the respondent their current total annual *household* income. Respondents are instructed to report income from all sources and include incomes from all adults in the household.

A.4 Baseline Data

The survey next asks respondents to choose the three most likely state and metropolitan areas where they may live 5 years from now. The first question is the most likely, the second is the second-most likely, and the third is the least likely.

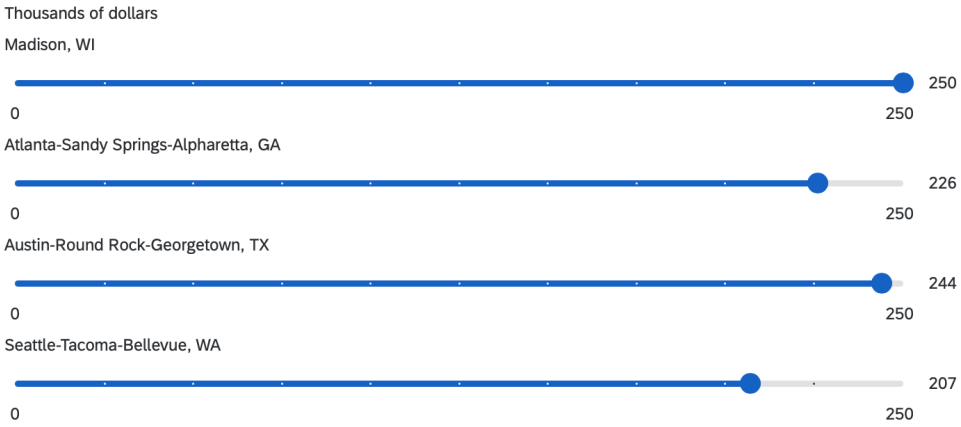
To help fix ideas, going forward we will consider an example where the most likely location is Atlanta, GA, the second most likely location is Austin, TX, the third most likely location is Seattle, WA. In this example, the respondent's hometown is Madison, WI and current location is Washington, DC.

After the respondent reports the three places he or she is most likely to live, the respondent is asked to report the probabilities he or she will be "living in these locations 5 years from now." The set of locations includes the respondent's hometown, and the respondent is asked to make sure the probabilities sum to 100 percent. A sample response is below.



We will call the collection of the three most likely locations respondents expect to live 5 years from now and the home town as the respondent's "likely future locations."

The respondent is then asked to report his or her expected annual total income in 5 years in each of the likely future locations. The respondent answers these questions by moving a slider, shown below.



The next questions ask about the presence of savings and friends and family in each of the likely future locations as well as expected future costs of living.

- We first ask respondents if they have enough cash on hand to "cover expenses for a few months in the event of a job loss or a relatively small event causing some unanticipated expenses." A sample response is shown below.



- We next ask if respondents have a social network of family or friends in each of the likely future locations. In each location, the answer is restricted to be "yes" or "no."

- The respondent is then asked, *In the future, if your life circumstances become very difficult, in which location(s) do you expect to get the most support from your social network of family and friends?* The respondent is asked to rank the likely future locations from 1 (most likely to get support) to 4 (least likely).

- Next, the respondent is asked, *Thinking about your hobbies and things you like to do for entertainment, where do you think you will have the most fun?* The respondent is asked to rank the likely future locations from 1 (most fun) to 4 (least fun).

- Finally, the respondent is asked, *Given the sorts of goods and services you like to buy and given the kind of housing unit you wish to live in, which locations do you think would be most expensive for you?* The respondent is asked to rank the likely future locations from 1 (most expensive) to 4 (least expensive).

A.5 Experimental Variation in Hypothetical Scenarios

The next two blocks of questions are key for estimating the marginal utility of consumption. First, we ask how probabilities over location chioces change as assumptions about income and the presence of family and friends vary. Instructions leading up to this block of questions are as follows:

> *Earlier you told us four places you could imagine living 5 years from now. You also gave us your best guess about what your annual income would be in each location, whether family would be present, and your probability of living in each place.*
>
> *Now we are interested in how the probability of living in these locations would be affected if your anticipated income or family ties were changed, holding all other aspects of the location the same.*
>
> *For your convenience, we will also show your reported expectation of income and presence or not of family/friends in parenthesis below the changed income and family ties.*

Respondents are shown new values of income and family and friends in each location, with the baseline values shown in parentheses. We select the new values quasi-randomly to span a large support of possible variation. The variation in incomes and presence of family and friends for each MSA, by scenario, is shown below

		Income Growth Factor			
		hometown	1	2	3
	1	1.2520	0.9162	1.3590	0.8913
	2	0.7849	1.1480	0.8662	1.2660
Scenario	3	1.2090	0.8617	0.6198	1.1050
	4	0.8848	1.4440	1.2390	0.7793
	5	0.7254	0.7256	1.3950	1.1870
	6	1.0650	1.2130	0.9453	0.7797
	7	0.8673	0.8899	0.7091	1.4020
	8	1.4440	0.9491	1.1100	1.2090

		Family and Friends			
		hometown	1	2	3
	1	0	1	1	0
	2	0	1	1	0
Scenario	3	1	0	0	1
	4	1	0	0	1
	5	1	0	0	1
	6	1	1	0	0
	7	0	0	1	0
	8	0	1	1	1

Continuing with our example, in a first scenario respondents are shown the following:

Please report your revised chances of living in each location if, in 5 years, your expected household income in these locations and presence or not of family/friends are different than what you expected as shown below, while holding all other aspects of the locations the same.

Location	Future Income	Family/ Friends
Madison, WI	313k (was 250k)	No (was Yes)
Atlanta-Sandy Springs-Alpharetta, GA	206k (was 226k)	Yes (was No)
Austin-Round Rock-Georgetown, TX	332k (was 244k)	Yes (was Yes)
Seattle-Tacoma-Bellevue, WA	184k (was 207k)	No (was No)

Given the new assumptions, respondents are asked to recompute the probabilities they live in each of the likely future locations in 5 years:[14]

*Your **reported % chances of living in each location is indicated in parentheses.***
*Please make sure that **your new % chances sum to 100**.*

Madison, WI (10%)	New%	15
Atlanta-Sandy Springs-Alpharetta, GA (20%)	New%	15
Austin-Round Rock-Georgetown, TX (40%)	New%	45
Seattle-Tacoma-Bellevue, WA (30%)	New%	25
Total	New%	100

Respondents are shown 9 scenarios in total: baseline plus 8. Call this entire set of 9 scenarios as the "probability scenarios" which help us identify the elasticity of location choice with respect to expected income and the presence of friends and family.

The next block of questions allows for estimation of the marginal utility of consumption in each of the likely future locations, as we show later. Instructions for this block of questions are as follows:

> *The following questions explore a circumstance in which your income unexpectedly and permanently jumps **after** you have chosen your future location. We will then ask the same two questions under a number of different scenarios where your expected household income and presence or not of family/friends might differ from your stated expectations.*

[14] Note the probabilities at the baseline assumptions are shown in parentheses.

1. The first question will ask you to choose the location in which you would benefit most from the permanent jump in income, had you chosen to live in that location. Call the benefit arising from this permanent jump in income in the location you choose your "change in happiness." Notes that this location may not necessarily be the one where you are most likely to move. Simply think of where the permanent jump in income would be most helpful had you chosen to live there.

2. Next, we will ask you to imagine living in the other locations 5 years from now, and how much of a jump in income would be needed for your change in happiness at each place to be equal to the change in happiness at the location in the first question.

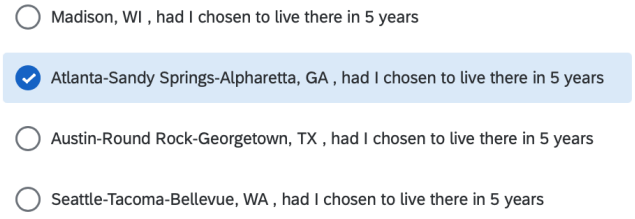
Respondents will be asked the same two questions nine times. The first time, they will be shown a scenario corresponding to their baseline assumptions. The next eight times, they will be shown scenarios that are the same as the scenarios shown in the probability questions. In each of these nine scenarios, respondents will be instructed to hold all factors other than income and the presence of family and friends the same.

- The first question is as follows: *Fast forward 5 years. Out of many possible future versions of yourself, think about 4 versions A, B, C, and D each of whom has chosen to live in one of the locations you have specified and with the incomes and presence of family/friends as given below.*

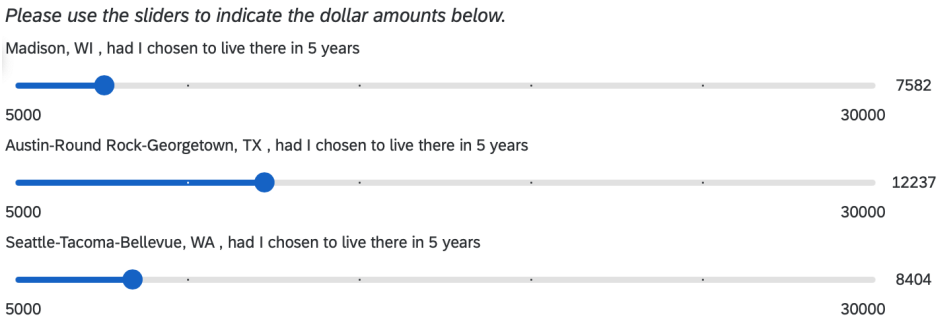
  In the first scenario, respondents are shown baseline assumptions

Location	Future Income	Family/ Friends
Madison, WI	250k	Yes
Atlanta-Sandy Springs-Alpharetta, GA	226k	No
Austin-Round Rock-Georgetown, TX	244k	Yes
Seattle-Tacoma-Bellevue, WA	207k	No

  The question continues: *Fast forward 5 years and imagine yourself in each of the locations below. In which location would you benefit the most from an unexpected and permanent increase of $5,000 in household income, had you chosen to live there? Note that this location may not necessarily be the one where you are most likely to move. Call the benefit arising from this permanent jump in income in the location you choose your "change in happiness."* Respondents choose a location from the list of likely future locations:



- The second question is: *Now, for each of the other versions living in the other locations, indicate the size of the jump in household income that they would need living in those locations 5 years from now such that their change in happiness after receiving this jump in household income is the same as the change in happiness from the $5,000 jump in household income given to the version living in [Atlanta-Sandy Springs-Alpharetta, GA].*[15] Respondents then adjust a slider for each of the other three likely future locations to answer that question:



The respondents then answer the same 2 questions 8 more times, but with the assumptions about income and friends and family in each location changed, for example:

Location	Future Income	Family/ Friends
Madison, WI	313k (was 250k)	No (was Yes)
Atlanta-Sandy Springs-Alpharetta, GA	206k (was 226k)	Yes (was No)
Austin-Round Rock-Georgetown, TX	332k (was 244k)	Yes (was Yes)
Seattle-Tacoma-Bellevue, WA	184k (was 207k)	No (was No)

A.5.1 Economic Policy and Political Leanings

After we ask these 18 questions (2 questions under 9 sets of assumptions), we then ask a block of 7 questions about economic policy and political leanings that we borrow from Kuziemko et al. (2015). We are interested in whether political preferences on views on redistribution across agents are correlated with expected variation in their own marginal utility of consumption across potential likely future places. We list these questions below in italics:

- *On economic policy matters, where do you see yourself on the liberal-conservative spectrum?*. Respondents choose from "Very Liberal," "Liberal," "Moderate," "Conservative," and "Very Conservative."

- *In general, how important do you think it is to stay informed about economic policy?"* Respondents choose from "Very important," "Somewhat important," "Not very important," and "Not Important at all."

[15] Note that the location in brackets will be the location chosen as the answer to the previous question.

- *What would you say are the main reasons why you wish to be well informed about economic policy?* Respondents choose one of the following responses "Affects personal finances," "Affects business or profession," "Relevant to stock market and investments," "Economic issues are important politically and might affect my vote," and "To be a responsible citizen, I like to keep informed."

- *How knowledgeable do you consider yourself on economic policies and issues?* Respondents choose one of the following responses "Highly knowledgeable," "Somewhat knowledgeable," "Not very knowledgeable," and "Not knowledgable at all"

- *Do you agree that the government should assist people in response to temporary crises, for example cash transfers to people that lost their job during a recession or households that lost their house and possessions due to hurricane, fire, tornado, or flood?* Respondents choose one of the following response, "Definitely," Probably," "Probably not," and "Definitely not."

- *Do you agree that the government should implement welfare programs using tax dollars?* Respondents choose one of the following response, "Definitely," Probably," "Probably not," and "Definitely not."

- *Your responses show that a little extra money in certain scenarios is useful. Do you believe that the government should help individuals with some cash assistance in such scenarios.* Respondents choose one of the following responses, "Definitely," Probably," "Probably not," and "Definitely not."

A.5.2 Demographic Questions

In the final 9 questions of the survey, we ask each respondent to list (1) his or her age (restricted to one of 7 intervals), (2) gender (male, female, or other), (3) if the respondent is of Hispanic, Latino, or Spanish origin (yes/no), (4) ethnicity or race (one of six labels including "Other,"), (5) current marital status, (6) number of adults in the household, (7) children under 18 in the household, (8) highest level of education (one of nine categories), and (6) employment status over the past 12 months (one of seven categories).

B Empirical Bayes Algorithm

Consider a collection of n parameter estimates $\{\hat{\theta}_i\}_{i=1}^n$ with associated standard errors $\{\hat{\sigma}_i\}_{i=1}^n$. The hierarchical model is specified as:

$$\begin{aligned} \text{Truth:} \quad \theta_i &\sim N(\mu, \tau^2) && \text{(Prior)} \\ \hat{\theta}_i|\theta_i &\sim N(\theta_i, \hat{\sigma}^2) && \text{(Sampling)} \end{aligned}$$

where:

- μ is the population mean, τ^2 is the prior variance
- θ_i represents the true parameter value for observation i
- $\hat{\sigma}^2$ is the estimated sampling variance. To estimate this, we:
  1. Estimate θ_i for each i separately using Canay (2011)
  2. Obtain residual vector for each i
  3. Stack residuals and estimate its variance to obtain $\hat{\sigma}^2$

Given the specified model, we employ an empirical Bayes approach where $\hat{\theta}_i \sim N(\mu, \hat{\sigma}^2 + \tau^2)$. This framework allows us to optimally combine individual estimates with information from the entire sample to reduce estimation error.

The log likelihood function for this specification can be expressed as:

$$\ell(\mu, \tau^2; \{\hat{\theta}_i, \hat{\sigma}^2\}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\hat{\sigma}^2 + \tau^2) - \frac{1}{2(\hat{\sigma}^2 + \tau^2)}\sum_{i=1}^{n}(\hat{\theta}_i - \mu)^2$$

Since the weights are all equal in our homoskedastic case, the estimate of the prior mean reduces to a simple average: $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_i$. Conditioning on this estimated mean, the likelihood function for τ^2 becomes:

$$\ell(\tau^2; \{\hat{\mu}, \hat{\theta}_i, \hat{\sigma}^2\}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\hat{\sigma}^2 + \tau^2) - \frac{1}{2(\hat{\sigma}^2 + \tau^2)}\sum_{i=1}^{n}(\hat{\theta}_i - \hat{\mu})^2$$

To estimate the variance component τ^2, we derive the first-order condition:

$$\frac{n}{2(\hat{\sigma}^2 + \hat{\tau}^2)} - \frac{1}{2(\hat{\sigma}^2 + \hat{\tau}^2)^2}\sum_{i=1}^{n}(\hat{\theta}_i - \hat{\mu})^2 = 0$$

Solving this equation yields the maximum likelihood estimate:

$$\hat{\tau}^2 = \frac{1}{n}\sum_{i=1}^{n}(\hat{\theta}_i - \hat{\mu})^2 - \hat{\sigma}^2 \tag{14}$$

This approach produces identical shrinkage factors for all observations, where $\lambda_i = \lambda = \frac{\hat{\tau}^2}{\hat{\sigma}^2+\hat{\tau}^2}$.

The resulting empirical Bayes estimate for each observation is a weighted average: $\tilde{\theta}_i = (1-\lambda)\hat{\mu} + \lambda\hat{\theta}_i$, reflecting the balance between individual estimates and the sample-wide mean in proportion to the estimated signal-to-noise ratio. The overall signal-to-noise ratio, identical for each observation i, can be expressed as $\frac{\hat{\tau}^2}{\hat{\sigma}^2} = \frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{\theta}_i-\hat{\mu})^2}{\hat{\sigma}^2} - 1$.