NEW YORKERS COMMUTE MORE EVERYWHERE: CONTRAST EFFECTS IN THE FIELD

Uri Simonsohn*

Abstract—Previous experimental research has shown that people's decisions can be influenced by options they have encountered in the past. This paper uses PSID data to study this phenomenon in the field, by observing how long people commute after moving between cities. It is found, as predicted, that (i) people choose longer commutes in a city they have just moved to, the longer the average commute was in the city they came from, and (ii) when they move again within the new city, they revise their commute length, countering the effect their origin city had on their initial decision.

I. Introduction

Is that more promising job worth commuting downtown everyday? Does residing in a better school district compensate the discomfort of commuting 10 more minutes to work? Does getting away from Manhattan's high rents justify the 45-minute train ride? To answer these questions, people must trade off the utility of the benefits they obtain by the longer commute, with the disutility of the extra minutes spent traveling on a daily basis.

Based on what is known in psychology as *contrast effects*—people's perceptions and judgments being affected by earlier experiences—this paper proposes that people behave as if the disutility they derive from commuting is influenced by commutes they have observed in the past. In particular, the longer the average commute was in a city a person moved from, the higher the willingness to endure long commutes in the new city will be. This prediction found support in a sample of intercity movers from the Panel Study of Income Dynamics (PSID).

The most plausible alternative explanation for this finding is that movers from cities with longer commutes have a (stable) higher tolerance for commuting, either because they self-selected into those cities based on their commuting tolerance, or because they developed such tolerance after commuting long distances in those cities.

The empirical section presents several tests showing that unobserved stable heterogeneity in commuting tolerance cannot fully account for the data. Perhaps the most important one of these results is that, consistent with contrast

effects but not with stable unobserved differences, the effect of the previous city quickly dissipates as people remain in their new city. A more accurate title for this paper might be "New Yorkers Commute More Everywhere—For a While."

One interpretation for this paper's findings is that people acquire commuting capital not only by "consuming" long commutes, as proposed by habit formation models, but also merely by living in a city where people commute a lot.

In what follows, section II briefly reviews relevant literature, section III links such literature to the specific realm of travel demand, section IV presents the empirical analysis, and section V concludes.

II. Background

Economists have, until very recently, typically left the task of studying factors that systematically affect people's preferences to other social scientists. In recent years, however, there has been growing interest in incorporating psychology's insights on preference formation into economics. For example, Rabin (1993) and Fehr and Schmidt (1999) discuss moderators of prosocial preferences, Laibson (2001) the effects of environmental cues on preferences, Loewenstein (1996) the role of emotional arousal in self-control, and Caplin and Leahy (2001) the importance of anticipatory feelings; Koszegi and Rabin (2004), Tversky and Kahneman (1991), and Sugden (2003) introduce models of reference-dependent preferences.

An interesting insight from psychology that has not yet made its way into economics is what are referred to as *context effects,* defined by McFadden (1999) as situations where choices are affected by the setting in which alternatives are offered. Huber, Payne, and Puto (1982), for example, found that adding an extreme alternative to a choice set can make other extreme, but not as extreme, alternatives appear to be "compromise" options, *increasing* their market share.

A context effect that is particularly relevant to this paper is *background contrasts,* which occur when previously faced options affect current decisions. Simonson and Tversky (1992) first documented this effect in an experiment where subjects made sets of two choices, the first being the experimental treatment. The manipulation consisted in having half the subjects make a choice between two options that

had an implicit *high* cost for an attribute, and the other half making a choice with an implicit *low* cost for the same attribute. The second choice was the same for all subjects. Consistent with the authors' hypothesis, subjects who were in the high-cost treatment were more likely to choose the expensive option in the second choice.

Contrast effects have also been found in nonhuman animals. Waite (2001) designed a study similar to that of Simonson and Tversky, using gray jays as subjects. All birds in the experiment had a choice of obtaining one raisin with low effort or three raisins with high effort. As in Simonson and Tversky (1992), half the subjects faced first an expensive choice (where the three raisins required much more effort than just one) and half a cheap choice (where three were just as hard to obtain as one). Consistent with a background contrast effect, birds that had experienced the expensive background were more likely to pursue the three raisins in the second stage than those who had experienced the cheap one. More recently, in a closely related project, George Loewenstein and I show that the amount of money people spend on rent in a city they have just moved to is influenced by how expensive housing is in the city they came from (Simonsohn & Loewenstein, 2006).

III.   Background Contrasts and Commuting

This paper studies background contrasts by examining how average commutes in cities where people have lived in the past affect commuting decisions after individuals have moved to a new city. Note that unlike habit formation models, contrast effects occur when previously observed options, rather than previous choices, affect current decisions.

A standard utility model can easily be modified to incorporate background contrasts by treating previously observed options as complementary (that is, marginal-utility-enhancing) with current consumption. This approach is commonly used to model external influences on preferences (see for example, Laibson, 2001; Becker & Murphy, 1988, 1993; and Stigler & Becker, 1977).

In particular, let the utility function be $U_t = U(I_t, C_t; C_t^*)$, where I_t is the consumption in period t, C_t is the commuting time in period t, and C_t^* represents the background used to assess current tradeoffs, corresponding in this case to previously observed commutes. C_t^* is indexed by time because the background an individual uses to judge options is proposed to be updated as new budget constraints for commuting are encountered.

Background contrasts imply that the marginal rate of substitution between commuting and consumption increases after having experienced a high background, and hence we can characterize the chosen commute length at period t by individual i with the following reduced-form equation:

$$C_{i,t} = \beta_0 + \beta_1 X_{i,t} + \beta_2 C_{i,t}^*, \qquad (1)$$

where $X_{i,t}$ are the relevant characteristics of individual i in time period t. Because $C_{i,t}^*$ is not observable, we must use a proxy to estimate equation (1). In particular, we use average commute in the city where a mover came from ($\bar{C}_{i,t-1}$),

$$C_{i,t} = \beta_0 + \beta_1 X_{i,t} + \beta_2 \bar{C}_{i,t-1}, \qquad (1')$$

leading to

Prediction 1.   The average commute length in the city a mover comes from has a positive effect on the chosen commute in the new city.

As the mover remains in the new city, however, the background used for judging commuting options (C^*) will converge toward the new local reality, leading in turn to a change in the desired commute length. To assess the effect of this adjustment we subtract equation (1) evaluated at period t from it evaluated at period $t+1$, obtaining

$$\Delta C_{i,t+1} = \beta_1 \Delta X_{i,t+1} + \beta_2(C_{i,t+1}^* - C_{i,t}^*). \qquad (2)$$

Using again $\bar{C}_{i,k}$ to represent the average commute in the city where person i lives in period k, and assuming that after living in the new city for a year people fully adjust to their new background (that is, that $C_{i,t+1}^* = \bar{C}_{i,t}$), we get

$$\Delta C_{i,t+1} = \beta_1 \Delta X_{i,t+1} + \beta_2(\bar{C}_{i,t} - \bar{C}_{i,t-1}), \qquad (2')$$

leading to

Prediction 2.   As people stay in their new city, they will readjust their commute, countering the effect their previous city had on their initial choice.

The intuition behind prediction 2 is the following: a person who moves from New York to Pittsburgh, say, may initially choose to live 35 minutes away from work (well above the average commute in Pittsburgh). As he stays in Pittsburgh, however, the 35-minute trip will be evaluated as too long, eventually leading him to move closer to work.

Because prediction 2 deals with changes in behavior, all stable differences across individuals are controlled for; if people who come from cities with higher $\bar{C}_{i,t-1}$ commute more in period t because they are (in some unobserved way) different from those coming from lower $\bar{C}_{i,t-1}$, we would not expect them to systematically revise their commutes as they stay in their destination cities. This second prediction will hence be useful in ruling out alternative explanations, particularly those relying on stable unobserved differences across movers from different cities.

Equations (1′) and (2′) highlight that for movers between cities there are two sources of variation in C^*. One comes from differences in previously observed commutes among individuals arriving from different cities in period t; the second comes from the updating of C^* within individuals as they remain in the new city. Each of these two sources of variation in C^* leads to a specific prediction, both derived from the hypothesis that, in equation (1), $\beta_2 > 0$. The data

TABLE 1.—COMPARISON OF REASONS FOR MOVING BETWEEN CITIES AND WITHIN CITIES

Reason for Moving	(1) Between Different Cities	(2) Within Same City
Job-related	36.0%	6.8%
Live closer to work	9.9%	14.6%
Housing-related (increase)	10.9%	18.2%
Housing-related (decrease)	5.5%	8.1%
Become owner or got married	10.5%	21.8%
Neighborhood, or closer to family or friends	3.9%	3.8%
Outside events (evicted, divorce, job transfer)	12.8%	13.2%
Mixed reasons	8.2%	10.1%
Don't know or refuse to answer	2.4%	3.6%
Total	100.0%	100.0%

then lend themselves to estimating β_2 through two independent regressions. As we shall see, the two methods lead to similar estimates of β_2.

IV. Empirical Analysis

A. Source of the Data

To study the influence of city-level conditions on individual decisions, it is necessary to combine individual and aggregate data. For the former, the PSID is ideal. It includes yearly data on a wide variety of economic decisions made by individuals, including city of residence and the length of their commute to work (in both distance and time).[1]

Although the analysis of interest deals with the behavior of movers between cities, the empirical analysis could be conducted using all households in the PSID, adding the relevant slope dummies for households that relocated to new cities; alternatively the analysis could be conducted exclusively on these intercity movers. Table 1 reports the distribution of answers to the question "Why did you move?" for both types of movers. Between- and within-city movers report significantly different motives for having moved, suggesting their moving decision should be analyzed separately. The analysis, therefore is conducted only on movers between cities, excluding people who did not move or who moved within a city.

The PSID collected data on commuting until 1986. The individual-level data, then, consist of PSID households that moved between two different MSAs in the 1972-to-1986 period.[2] Observations where commuting time was reported to be less than 2 minutes or more than 90 (one-way) were excluded from the analyses.[3]

For the city-level data, the PSID is not an adequate source, as it contains less than 5 observations for most MSAs. Instead, I used the Journey-to-Work Supplement of the 1980 Census.[4] In particular, I calculated the average commute length and proportion of people using different transportation modes in every metropolitan statistical area (MSA) in the United States, and then matched these figures to the PSID using the geographic data file referred to in footnote 1. For each household, the year prior to the move is referred to as $t-1$, the year of the move as t, and the year following it as $t+1$.

B. Data Description

The main reason for focusing on movers between cities is to observe decisions made by individuals who currently face commuting options that differ from those they have faced in the past (that is, individuals who face a *contrast* between current and *background* conditions).

Cities have an average commute (for drivers) of 21 minutes, with a standard deviation—across cities—of 2.9 minutes, or 13.8% of the mean. The average move in the sample occurred between cities that differ in their average commute by 13.6%; this figure indicates that cities between which individuals are moving indeed differ in their commuting conditions.

An important feature of the sample is that the moves occurred between a wide array of cities (204 origin cities and 218 destination cities). Of all moves in the sample, 83% were between a unique origin-destination city pair, and 97% of them were between city pairs that are repeated a maximum of three times.

C. Empirical Analyses

Testing Prediction 1: According to prediction 1, the average commute length in the city a mover comes from has a positive effect on the chosen commute in the new city. This section presents regression analyses that support this prediction. Alternative explanations for the findings, includ-

---

[1] Some of the data used in the analysis are derived from the Sensitive Data Files of the PSID, obtained under special contractual arrangements designed to protect the anonymity of respondents. These data are not available from the author. Persons interested in obtaining PSID Sensitive Data Files should contact PSIDHelp@isr.umich.edu.

[2] In 1982 the PSID didn't collect data on commuting time, so that year is not included in the sample.

[3] These corresponded to the 1st and 99th percentiles of observations, respectively. The fact that many of the respondents reporting commuting times of less than 2 minutes report distances from work superior to 10 miles, and many of those with commutes longer than 90 minutes report distances of less than 5 miles, validates the presumption that these observations correspond to errors.

[4] The Journey-to-Work Supplement is available at the Inter-University Consortium for Political and Social Research (ICPSR) Web site (http://www.icpsr.umich.edu).

TABLE 2.—TRAVEL DEMAND REGRESSIONS ESTIMATES

			Dependent variable: $C_{i,t}$			
	(1)	(2) Adds Avg. Commute City of Origin	(3) Adds Year Fixed Effects	(4) Adds own commute $t-1$	(5) Adds own commute $t-2$	(6) Only car commuters
Independent variable	Baseline					
Constant	−7.664	−18.889	−19.950	−16.776	−22.971	−9.406
	(7.703)	(8.691)	(8.996)	(8.733)	(9.050)	(8.072)
Family income (in $1,000s)	0.048	0.043	0.038	0.034	0.046	0.051
	(0.021)	(0.021)	(0.021)	(0.021)	(0.024)	(0.020)
Number of adults in household	−2.009	−1.761	−1.816	−1.512	−0.866	−0.879
	(1.119)	(1.116)	(1.125)	(1.108)	(1.195)	(1.019)
Number of children in household	1.240	1.187	1.182	1.203	0.835	1.421
	(0.578)	(0.580)	(0.568)	(0.566)	(0.619)	(0.037)
Age of head	0.380	0.402	0.380	0.302	0.378	0.176
	(0.369)	(0.363)	(0.365)	(0.355)	(0.394)	(0.393)
Age of head squared	−0.002	−0.003	−0.003	−0.002	−0.003	−0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
College dummy (1 if attended)	−0.391	−0.417	−0.046	−0.076	−0.698	−0.920
	(0.972)	(0.962)	(1.112)	(1.086)	(1.147)	(1.154)
Average commute in destination	1.025	0.852	0.825	0.819	0.744	0.603
	(0.174)	(0.174)	(0.173)	(0.171)	(0.192)	(0.183)
Average commute in city of origin	—	0.684	0.747	0.437	0.512	0.420
	—	(0.179)	(0.181)	(0.183)	(0.199)	(0.173)
Own commute in city of origin ($t-1$)	—	—	—	0.231	0.255	0.166
				(0.039)	(0.049)	(0.033)
Own commute in city of origin ($t-2$)	—	—	—	—	0.055	—
					(0.040)	
1/(Mills ratio)	—	—	—	—	—	0.023
						(3.162)
Year fixed effects	No	No	Yes	Yes	Yes	Yes
R-square	0.08	0.09	0.10	0.15	0.19	0.11
Number of observations	1,067	1,067	1,067	1,067	841	918

Robust standard errors below parameter estimates.

ing unobservable taste differences and imperfect information, are discussed later.

To test prediction 1, the following version of equation (1′) was estimated:

$$C_{i,t} = \beta_0 + \beta_{1,1}\bar{C}_{i,t} + \beta_{12}C_{i,t-1} + X_{it}\beta_{1,3} + \beta_2\bar{C}_{i,t-1}, \quad (1'')$$

where $X_{i,t}$ includes the number of adults and children, family income, age of household head, and a dummy for attending college. $\bar{C}_{i,t}$ is included to control for variation in commuting conditions across different cities. Table 2 reports the estimates for several specifications of equation (1″). Column 1 is the baseline specification, column 2 incorporates $\bar{C}_{i,t-1}$, column 3 adds yearly fixed effects, column 4 adds $C_{i,t-1}$, column 5 adds $C_{i,t-2}$, and column 6 restricts the analysis to drivers.

Most of the control variables' parameter estimates are hard to interpret, as these variables affect the desired commute length through diverse, often opposing mechanisms. For example, the number of children in a household may increase the desired commute length by restricting the school districts an individual is willing to move to (for example, a parent that would otherwise commute 10 minutes chooses to live farther from work in order to enjoy a

better school district). On the other hand, children may increase the opportunity cost of time and hence reduce the desired commute. Likewise, higher income may increase the opportunity cost of time (decreasing the desired commute length) but may also be the result of compensatory wages for longer commutes. Similar reasoning applies to the number of adults in the household, education, and age; their estimates, therefore, are not economically meaningful.

Column 2 incorporates the key variable of interest: $\bar{C}_{i,t-1}$. Its effect on $C_{i,t}$ is estimated as positive and significant ($p < 0.01$). Consistent with prediction 1, individuals coming from cities with longer average commutes choose to commute significantly longer in their new city than their peers coming from cities with shorter commutes. In column 3 yearly fixed effects are added to the regression, and they do not have a significant effect on any of the coefficient estimates.

Column 4 includes the individuals' own commute in the previous city ($C_{i,t-1}$) as a control. Previous commuting choices are likely to be predictive of current choices for both psychological reasons (for example, by acting as a reference point) and traditional economic reasons (unobserved heterogeneity is captured by previous choices). It is

therefore not surprising that $C_{i,t-1}$ is a significant predictor of $\bar{C}_{i,t}$. What's interesting is that $\bar{C}_{i,t-1}$ remains positive and significant ($p = 0.017$) after its inclusion. As we shall see later, this will be useful for ruling out unobserved heterogeneity as an alternative explanation.

In light of the fact that one single decision may be an unreliable measure of taste or previous commuting experience, column 5 includes the individual's commute of the previous two years as controls. $C_{i,t-1}$ can be a noisy measure of taste for two reasons. (i) classical measurement error (individuals misreport the actual commute length, or the PSID makes a mistake in recordkeeping), and (ii) individuals may have not been commuting the optimal distance in the past, so that the actual commute is a noisy measure of the desired commute. Adding the 2-year lag should reduce both sources of noise.

If noise in $C_{i,t-1}$ was the reason why $\bar{C}_{i,t-1}$ remained significant after its inclusion, reducing such noise should take away power from the estimated coefficient of $\bar{C}_{i,t-1}$. On the contrary, the point estimate of $\bar{C}_{i,t-1}$ is slightly strengthened by the inclusion of a cleaner control of previous choices. Note that because $C_{i,t-1}$ and $C_{i,t-2}$ are strongly correlated ($r = 0.42$), their individual significance tests should be interpreted with caution.

Commuting times are dramatically different for drivers and users of public transportation (in this sample, 22 and 38 minutes, respectively). It is possible therefore, that the effect of $\bar{C}_{i,t-1}$ on $C_{i,t}$ operates (exclusively) through the chosen mode of transportation, that is, that the origin city affects only willingness to use public transportation, and not the commute length per se. To assess this possibility, column 6 presents the results of estimating equation (3) only on drivers. The parameter estimate for $\bar{C}_{i,t-1}$ is still positive and significant, and of a similar magnitude to the one found for the whole sample, suggesting that the effect of C^* doesn't operate exclusively through the chosen mode of transportation. Because this regression was estimated on a restricted subsample (excluding nondrivers), a two-step Heckman procedure was employed. The results from the first stage (not reported) show that individuals who drove in the previous city, moved to cities with higher proportions of drivers, or had higher incomes were more likely to drive in their destination city.[5]

Testing Prediction 2: According to prediction 2, as individuals stay in their new city, C^* adjusts to local commuting standards and the desired commute follows suit. According to prediction 2, then, if an individual has moved from a long-commute city to a short-commute city ($\bar{C}_{i,t} <$ $\bar{C}_{i,t-1}$), the individual will initially commute too much and will choose a shorter commute in the future ($C_{i,t+1} < C_{i,t}$).

Because prediction 2 deals with changes in behavior, all stable differences across individuals are controlled for: if people who come from cities with higher $\bar{C}_{i,t-1}$ commute more in period t because they are (in some unobserved way) different from those coming from lower $\bar{C}_{i,t-1}$, we would not expect them to systematically revise their commutes as they stay in their destination cities. If, on the other hand, they are only temporarily affected by a contrast between current and past realities, once this contrast fades away, so should its effect on desired commute.

To test prediction 2, a regression was estimated with adjustment of commuting time between t and $t + 1$ as the dependent variable, and difference in average commuting time between current and previous city as the key predictor. This regression corresponds to equation (2′) in the previous section. Recall that the β_2 that measures the influence of the difference in average commutes between cities in equation (2′) is the same β_2 estimated for equation (1″).

We must restrict the analysis to people who move again within a year of arriving at the new city, because only for them is the change in desired commute length observable. The results are presented in Table 3.[6]

Column 1 shows the results without any covariates, column 2 adds changes in observables (income, number of adults in the household, and number of children), and column 3 presents the results of a two-stage Heckman procedure controlling for the potential selection bias arising from conducting the analysis on the subsample of people who moved again (first stage not reported).

As predicted, β_2 is estimated as positive and significant: as movers stay in their new city, they systematically readjust their commuting length, countering the initial effect that the previous city had on their choice in the new city. The point estimate of β_2 (around 0.63) is similar to the estimate of β_2 obtained in the testing of prediction 1. This suggests that after a year in the new city, those that move again completely reverse the initial effect that their previous city had on how far they live from work.[7]

D. Alternative Explanations

The previous subsections show evidence consistent with predictions 1 and 2: the longer the average commute in the city of origin, the longer the commute a newly arrived individual chooses in the destination city; but as she stays in the new city and moves again, her commute is revised, countering the initial effect of previously observed com-

---

[5] The second stage incorporates the Mills ratio, which is a function of other independent variables. In order for this stage to be identified beyond nonlinearities, it is necessary for the first stage to include variables which are excluded in the second stage. This is clearly the case here, where almost all variables from the first stage are excluded from the second.

[6] If the regression is estimated using the whole sample, the qualitative nature of the results does not change. The point estimate is reduced, however, as a result of incorporating many observations where instead of the change in desired commute length we observe a 0. The coefficient is still significant at the 5% level.

[7] Indeed, if equation (1″) is estimated using the subsample of re-movers and with commute in $t + 1$ as the dependent variable, $\bar{C}_{i,t-1}$ is no longer a significant predictor.

TABLE 3.—ADJUSTMENT OF COMMUTE LENGTH IN $t + 1$

Independent Variable	Dependent variable: $\Delta C_{i,t+1} = C_{i,t+1} - C_{i,t}$		
	(1) Baseline	(2) Adds Controls	(3) 2-Stage Heckman
Intercept	−0.502	−0.519	0.733
	(0.722)	(0.750)	(41.40)
$\bar{C}_{i,t} - \bar{C}_{i,t-1}$	0.631	0.627	0.627
	(0.188)	(0.189)	(0.201)
Change in family income	—	−0.019	−0.019
	—	(0.027)	(0.028)
Change in number of adults	—	1.663	1.499
	—	(2.017)	(5.89)
Change in number of children	—	−0.268	−0.220
	—	(1.810)	(2.29)
R-square	0.03	0.03	0.03
Number of observations	340	340	340

Robust standard errors below parameter estimates.

mutes. These two findings are consistent with the notion of contrast effects, but they may also be consistent with other explanations. In this subsection three specific alternative explanations are discussed: habit formation, self-selection into cities, and imperfect information.

Stable Taste Differences: In the presence of habit formation, past consumption increase the marginal utility of future consumption. If people's tolerance for commuting increases with "consumption" of long commutes, then movers coming from cities with higher $\bar{C}_{i,t-1}$ might choose longer commutes in the new city because they have commuted more (on average) in the past, and have developed a habit of it (they have accumulated more *commuting capital*).[8]

A self-selection-based explanation reverses causality between taste and past behavior: if people incorporated commuting considerations in their choice of what city to live in, those who dislike commuting the least would be more likely to have lived in cities with longer commutes. Movers from cities with higher $\bar{C}_{i,t-1}$ would commute more in the new city, not because they use the past as a yardstick to assess the present, but rather, because they have never been too much bothered by long commutes.

Because of taste formation and self-selection, then, movers coming from cities with different commuting conditions may differ in their idiosyncratic taste for commuting. In what follows I discuss four different approaches to assess the plausibility of this alternative explanation. All four suggest that unobserved heterogeneity is not sufficient to explain the results presented in the previous section.

(i) Inclusion of previous choices. If $\bar{C}_{i,t-1}$ correlates with $C_{i,t}$ because it is proxying for previous commuting choices, and hence with unobserved taste, then once

previous choices are directly controlled for, the effect should be eliminated. Note that, even in the absence of taste formation or self-selection, the significance of $\bar{C}_{i,t-1}$ is expected to drop after including $C_{i,t-1}$, due to collinearity. Finding an effect of $C_{i,t-1}$ on $C_{i,t}$ after such inclusion is hence a conservative test. As was shown earlier, including $C_{i,t-1}$ did not eliminate the effect of $\bar{C}_{i,t-1}$. Perhaps more importantly, adding $C_{i,t-2}$ had virtually no impact on the coefficient of $\bar{C}_{i,t-1}$, suggesting that the reason why $\bar{C}_{i,t-1}$ remains significant is not measurement error on $C_{i,t-1}$.

(ii) Excluding observable heterogeneity. Another way to asses the extent to which the estimated effect of $\bar{C}_{i,t-1}$ may be picking up *unobserved* heterogeneity is to evaluate how much variation in *observed* heterogeneity it picks up. Because there is no reason to suspect that the relationship between $\bar{C}_{i,t-1}$ and observables is any stronger than that with unobservables, a substantial change in the estimated coefficient of $\bar{C}_{i,t-1}$ after omitting all observables would be worrisome. To do this comparison, table 4 replicates the regression analysis presented in columns (4) and (5) of table 2, excluding all observable covariates. Comparing the point estimates for $\bar{C}_{i,t-1}$ of the two tables, we see that the exclusion of all covariates has an effect of around 10% on the point estimates (from 0.437 to 0.468 in one specification and from 0.512 to 0.573 in the other). Given that omitting (all) controls for observable heterogeneity wouldn't dramatically affect the estimate of $\bar{C}_{i,t-1}$, there is little reason to suspect that the omission of unobservable heterogeneity is what is driving the effect.

(iii) Aggregate commutes in t don't correlate with individual commutes in t−1. A sorting story proposes that individuals who chose to live in cities with long commutes dislike commuting less than individuals who chose cities with short commutes. This story would predict that individuals *currently* living in a city with long commutes do not

---

[8] Similarly, people may have invested in technology that lowers the disutility of commuting, for example, by purchasing large collections of books on tapes.

TABLE 4.—COMMUTING REGRESSION EXCLUDING OBSERVABLE HETEROGENEITY

	Dependent Variable: $C_{i,t-1}$	
Independent variable	Controlling for Previous Choice	Controlling for Two Previous Choices
Intercept	−10.185	−12.759
	(4.813)	(4.791)
Average commute in destination city	0.844	0.734
	(0.176)	(0.189)
Average commute in origin city	0.468	0.573
	(0.184)	(0.197)
Own commute in $t-1$	0.237	0.249
	(0.041)	(0.047)
Own commute in $t-2$	—	0.083
		(0.039)
Number of observations	1067	841
R-square	0.10	0.14

Robust standard errors below parameter estimates.

dislike commuting as much as individuals *currently* living in cities with short commutes; the relationship between aggregate and individual behavior, therefore, should also be found working backward in time (that is, people currently living in cities with long commutes should have commuted longer—ceteris paribus—in their origin city).

If, on the other hand, average commutes in the previous city have a *causal* effect on the present, as proposed in this paper, then $\bar{C}_{i,t}$ should not be correlated with $C_{i,t-1}$, because individuals choosing their commute in $t-1$ were not affected by the yet unseen background of commutes in their destination city. Table 5 reports the results of conducting this *reversed* regression. The dependent variable is the individual commute in the $t-1$, and the key predictor is the average commute in the destination city. The results show no association between average commutes in the destination city and individual-level commutes in the origin city, obtaining point estimates close to 0 and p-values far from standard significance levels ($p > 0.9$ in column 1, and $p > 0.3$ in column 2).

(iv) Stable differences are inconsistent with prediction 2. Finally, if people arriving from different cities have stable

TABLE 5.—CITY-LEVEL COMMUTES IN t DO NOT CORRELATE WITH INDIVIDUAL BEHAVIOR IN $t-1$

	Dependent Variable: $C_{i,t-1}$	
	(1)	(2) Adds own commute in
Independent Variable	Baseline	destination city
Constant	−17.956	−14.304
	(7.215)	(6.920)
Family income (in $1,000s)	0.003	−0.006
	(0.017)	(0.017)
Number of adults in household	1.310	1.993
	(0.989)	(0.975)
Number of children in household	0.035	−0.229
	(0.392)	(0.392)
Age of head	0.289	0.188
	(0.288)	(0.281)
Age of head squared	−0.002	−0.001
	(0.004)	(0.004)
College dummy (1 if attended)	1.301	1.402
	(0.957)	(0.934)
Average commute in destination city (t)	0.013	−0.163
	(0.169)	(0.170)
Average commute in city of origin ($t-1$)	1.410	1.274
	(0.159)	(0.153)
Own commute in destination city (t)	—	0.206
	—	0.036
R-square	8.93%	13.16%
Number of observations	1,067	1,067

Robust standard errors below parameter estimates.

taste differences that lead them to commute differently in their new city, then they should continue to make different choices in the future. If, on the other hand, their differences are short-lived and caused simply by the different backgrounds they have experienced, their behavior should coverage once they all share the same background.

The evidence presented earlier supporting prediction 2, therefore, is inconsistent with an explanation based on stable taste differences.

In sum, all four approaches fail to find evidence consistent with the notion that unobserved heterogeneity in taste, correlated with city of origin, may be behind the effect.

Imperfect Information: Imperfect information by itself would not lead to the reported effect of city of origin on commute length; if people have noisy estimates of what the commute tradeoffs are in their new city, they should choose suboptimally, but not systematically so.

Imperfect information plus a few auxiliary assumptions, however, could account for the reported pattern. For example, if one assumes that movers (1) are unaware of the commuting conditions in the new city, (2) organize their house (or job) search by distance from work (or home), and (3) have a stopping rule (that is, a reservation distance they are willing to travel), then movers from cities with shorter commutes would end up living closer to work because their stopping rule would be triggered earlier.

Such a process is in fact very similar to the interpretation of the data proposed in this paper, where the same objective commute is *perceived* differently (above or below the stopping rule) based on how long commutes have been in the past. Why would movers from a city with long commutes settle for a neighborhood that's farther from work and discontinue the search earlier than people moving from a city with short commutes? Perhaps precisely because their perception of a *reasonable* commute is affected by their experience.

Regardless of the specific auxiliary assumptions, an imperfect-information-based explanation is problematic. Because of the importance of the decision and the low cost of the information involved, it seems implausible to suggest that optimally obtained information may be driving the effect. If households search for information optimally, they should do so until the expected payoff of the information is equal to its cost. Information about commuting is very cheap—arguably almost unavoidably acquired by movers as they are looking for homes (or jobs) and as they travel around their new city—whereas the benefits of choosing the appropriate commute are sizable.[9] Note that I am not suggesting that newcomers will

have perfect information about commuting options in their new city, but only that they should be able to make *unbiased* estimates of the typical commute length.

V.   Conclusions

This paper documents that the average commute in the city movers came from has a positive influence on the commute they choose in the new city, consistent with the notion of contrast effects. As these movers stay in their new cities and make new commuting decisions, they counter the initial influence of their origin, again consistent with contrast effects, but not with alternative explanations that rely on imperfect information, habit formation, or selection into cities based on stable taste.

The notion that preferences are affected by previous consumption is by now widely accepted in economics. This paper suggests that, in addition, people behave as if previously observed options also influenced their current marginal utility. This effect appears to be short-lived, probably because new options replace previous ones in their role as background against which options are evaluated.

These findings have potentially important implications for empirical investigations of preferences. Applied economists from various fields use observed behavior to infer preferences in an array of different domains, such as wage differentials in labor economics, hedonic prices in housing economics, and travel costs in environmental economics, among many others. The validity of such studies relies on the assumption that the preferences that are revealed for the goods and services the researchers are studying are stable. If, as the results from this paper suggest, preferences are affected by the options they have encountered, future empirical work should strive to identify them and include them in the analysis, particularly if they are conducting the study in order to assess the impact of a specific policy change that may change the very conditions that are currently affecting preferences.

---

[9] With back-of-the-envelope calculations we can ballpark the effect size of C^* at around \$145 a year: The estimated coefficient of $\bar{C}_{i,t}$ is 0.6. The standard deviation of $\bar{C}_{i,t}$ is 2.9 minutes. With 10 trips per week and 50 weeks a year at a value of \$10 an hour we get [0.6 × (2.9 × 10)/60] × 50 × \$10 = \$145.

REFERENCES

Becker, G. S., and K. M. Murphy, "A Theory of Rational Addiction," *Journal of Political Economy* 96 (1988), 675–700.
——— "A Simple Theory of Advertising as a Good or Bad," *Quarterly Journal of Economics* 108 (1993), 941–964.
Caplin, A., and J. Leahy, "Psychological Expected Utility Theory and Anticipatory Feelings," *Quarterly Journal of Economics* 116 (2001), 55–79.
Fehr, E., and K. M. Schmidt, "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics* 114 (1999), 817–868.
Huber, J., J. W. Payne, and C. Puto, "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research* 9 (1982), 90–98.
Koszegi, B., and M. Rabin, "A Model of Reference-Dependent Preferences," University of California, Berkeley working paper (2004).

Laibson, D., "A Cue-Theory of Consumption," *Quarterly Journal of Economics* 116 (2001), 81–119.

Loewenstein, G., "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes* 65 (1996), 272–292.

McFadden, D., "Rationality for Economists?" *Journal of Risk & Uncertainty* 19 (1999), 73–105.

Rabin, M., "Incorporating Fairness into Game-Theory and Economics," *American Economic Review* 83 (1993), 1281–1302.

Simonsohn, U., and G. Loewenstein, "Mistake #37: The Impact of Previously Observed Prices on Housing Demand," *The Economic Journal* (2006), 82–106.

Simonson, I., and A. Tversky, "Choice in Context—Tradeoff Contrast and Extremeness Aversion," *Journal of Marketing Research* 29 (1992), 281–295.

Stigler, G. J., and G. S. Becker, "De Gustibus Non Est Disputandum," *American Economic Review* 67 (1977), 76–90.

Sugden, R., "Reference-Dependent Subjective Expected Utility," *Journal of Economic Theory* 111 (2003), 172–191.

Tversky, A., and D. Kahneman, "Loss Aversion in Riskless Choice—a Reference-Dependent Model," *Quarterly Journal of Economics* 106 (1991), 1039–1061.

Waite, T. A., "Background Context and Decision Making in Hoarding Gray Jays," *Behavioral Ecology* 12 (2001), 318–324.