Commute duration and health: Empirical evidence from Brazil

Rodrigo Oliveira [a], Klebson Moura [a], Jorge Viana [a], Robson Tigre [b,a], Breno Sampaio [a,*]

[a] Department of Economics, Universidade Federal de Pernambuco, Brazil
[b] Department of Economics, University of Bologna, Italy

A R T I C L E   I N F O

A B S T R A C T

There have been many empirical studies associating commuting time and health outcomes in the last few decades. Their general conclusion is that commuting and health are negatively related. The validity of their findings, however, is questionable, given their lack of good identification strategies to correctly account for omitted variables. In this paper, we analyze this relationship using a large and unique nationally representative sample of Brazilian individuals, coupled with the use of propensity score matching techniques, and the application of an exhaustive set of standard falsification tests and sensitivity analysis that may prevent one from claiming a causal link between the two variables. Our results indicate that individuals with more than one hour of commuting appear to have statistically higher probability, ranging from 1.9 to 4.6 percentage points, of reporting bad health status when compared to a person whose commuting time is less than one hour.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Increasing traffic intensity within urban areas, which is commonly pointed as a major issue in large cities, arises when the transport infrastructure does not follow changes in individuals' commuting dynamics. In Brazil this question gained much attention recently, given the urbanization rate has doubled in just a few decades and there was no significant improvement in transport infrastructure (Renkow and Hoover, 2000; Sohn, 2005). Also, more recent counter-cyclical economic policies (i.e. tax reduction on industrialized products such as automobiles and motorcycles) have favored a drastic increase in Brazilian automotive fleet, notably leading to consequences to commuting time.

Longer commuting is associated with higher monetary and opportunity costs incorporated in transportation. Those costs are even higher in urban centers lacking of good transportation infrastructure, as it seems to be the case for many Brazilian cities. The additional time spent moving from one place to another could otherwise be allocated to productive activities or devoted to leisure and physical activities. Either by reducing leisure and sleep time or by increasing the stress caused by traffic, most scientific estimates agree that longer commuting negatively affect a wide range of health indicators, such as perceived stress levels (Gottholmseder et al., 2009).

In addition, a not so recent trend of sub-urbanization has been observed in Brazil, a phenomena defined as the establishment of densely populated neighborhoods in peripheral regions of urban cities. This trend has been identified in the United States and in Europe (Frumkin et al., 2004; Sandow, 2011), and within some years could increase significantly the duration of commuting if the same pattern observed in those regions could be applied to Brazil.

Despite the scarce evidence in Brazilian literature, many studies claim to have identified a negative relationship between commuting duration and levels of mental and physical health, as is the case of the seminal paper of Costa et al. (1988). Novaco and Collier (1994) found commuting stress to be significantly associated with duration of commute using a

---

representative sample of 2591 southern California automobile drivers. Moczulski et al. (2007), using a sample of university students, found that students with a long drive to the university were more likely to suffer from overweight than those with short drives. Hansson et al. (2011), using a cross-sectional population survey in southern Sweden, identified negative associations between commuting duration and variables such as self-assessed health, sleep quality, mental health and work absenteeism. Karlström and Isacsson (2009) identified a weak effect of commuting on the use of health insurance, being stronger in low-income women. Nelen (2012) analyzed both subjective health outcomes, like health satisfaction and health status, and objective measures, like the number of doctor visits and whether or not someone has health problems, and concluded that commuting time was negatively related with health satisfaction and health status for Germany and additionally linked to larger number of doctor visits and a higher probability of being involved in an accident in the UK. Finally, using a quasi-experimental approach, Wener et al. (2003) demonstrates that improvements in urban transportation could mitigate these impacts on well-being by showing that the creation of a new direct route between two U.S. locations improved health indicators of workers who started to use the new route instead of the former.

Most of the papers in this literature, however, do not recognize the problems incurred when selection into treatment is determined jointly by variables that are observed and unobserved by the researcher. In other words, given that individuals are not randomly assigned to live near or far from work, even when accounting for important observable differences between individuals, such as income and education, it is still hard to argue that other unobserved variables affecting health outcomes are uncorrelated with the likelihood of having a short commuting time. Therefore, one may question the validity of the findings presented so far, given their lack of good identification strategies to correctly account for the well known omitted variable bias.

Concerned about the lack of robust works about commuting, specially considering the case of Brazil, as the country is already facing problems related to this subject, our main objective in this article is to empirically analyze the effect of longer commuting duration on subjective individual health measures, precisely defined as self-assessed health status. Based on the aforementioned references, our hypothesis is that longer commuting duration should lead to a higher probability of individuals' reporting their health as "not good."

To the best our knowledge our work is the only one among commuting studies to encompass at the same time the use of a large nationally representative sample, the application of an exhaustive set of standard falsification tests and also to put under scrutiny the possibility of omitted variable be driving our results. For that, we use four different empirical strategies. We first estimate the effect of longer commuting on health via a propensity score matching (PSM) estimator, to balance covariates and to obtain more reasonable estimates when compared to effects obtained from simple OLS. Secondly, we apply the simulation-based sensitivity analysis presented by Ichino et al. (2008) to check the robustness of our estimates under the failure of the Conditional Independence Assumption (CIA), required for the validity of results obtained via PSM. Third, we implement the two-stage estimator proposed by Lewbel (2012) that exploits heteroskedasticity to generate an instrumental variable (IV) for identification, a method now widely considered in empirical analysis when there are no exclusion restrictions available. Finally, we consider the procedure recently developed by Oster (2013, 2015) to estimate bounds for the treatment effect under the notion that movements in the coefficient of interest, when including and not including controls for which one is concerned about omitted variables, are informative about the remaining bias from unobservables.

Our results show that individuals with more than one hour of commuting have 1.9–4.6 percentage points increase in the probability of reporting a bad health when compared to a person whose commuting time is less than one hour. As part of our empirical implementation, we first show that our matching procedure successfully balanced observed characteristics. Before matching, long commuters differed substantially from short commuters given out of the 19 variables considered in our model, 16 variables had mean differences that were statistically different at conventional levels. After matching, none of the variables were statistically different for the two groups. Secondly, unobservables not accounted for in our model does not seem to be responsible for our results. We applied Ichino et al. (2008) falsification test to show that any unobserved factor correlated with each of the covariates used in this study would not be sufficient to drive our estimated average treatment effect to zero. Lewbel's (2012) method yields similar results when compared to those obtained via OLS and propensity score matching, namely the parameter of interest was 14% and 19% greater than OLS and PSM, respectively. Finally, results obtained via Oster's (2013, 2015) methodology imply that the effect of unobservable variables in explaining health differences related to longer commuting would have to be at least 2.857 times stronger than the effect of the observable variables in order to fully explain the negative effect of long-commuting on self-reported health, a very unlikely scenario.[1] The consistency of results across these four alternative empirical strategies lends additional credibility to our claim that we find the causal effect of commuting on self-reported health.

Although our results based on ordinary least squares and other identifying methods did not differ by much, which may induce one to conclude that unobserved variables are not the driving force behind our statistical findings, our study may guide future researchers in testing the robustness of their estimates under weaker assumptions regarding selection on unobservables, specially those using data for the US or European countries, whose patterns of urban mobility, for instance, differ by much from those observed in Brazil. A recent unique survey collected in one of the largest metropolitan regions of Brazil showed that proximity to work played a despicable role in household choice (Tigre et al., 2014). These country-level differences, therefore, may cause large shifts to naive OLS estimates leading to incorrect causal interpretations.

---

[1] This implies a lower and upper bound for the average treatment effect of −.019 and −.029, respectively.