ESSAYS IN APPLIED ECONOMICS, HOUSEHOLD BEHAVIOR, AND ENVIRONMENTAL ECONOMICS

WANG YI

(*B.S.(Hons), NTU*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECONOMICS

NATIONAL UNIVERSITY OF SINGAPORE

2017

Supervisor:
Associate Professor Alberto Salvo

Examiners:
Associate Professor Liu Haoming
Dr Tan Soo Jie Sheng
Assistant Professor Kim Seonghoon, Singapore Management University

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Signed:

Date: **10 May, 2017**

Acknowledgements

This thesis would have been impossible had it not been for the assistance of professors, friends, classmates and my family. I am grateful to all people that have helped me and made this thesis possible.

First of all, it is with immense gratitude that I acknowledge the guidance and support of my supervisor, Professor Alberto Salvo. His enthusiasm, patience, knowledge and inspiration for research have encouraged me and helped me when I was writing this thesis.

I would like to thank the rest of my thesis committee, Professor Yang Nan, Professor Yi Junjian, and Professor Aamir R. HASHMI, for their valuable comments and suggestions. I have benefited a lot from them, who are patient, encouraging and helpful.

I would like to express my very great appreciation to all the participants in SIG reading group and my graduate research seminars, who gave me lots of valuable comments.

I would also like to thank all my PhD classmates and friends, without whom I would have never gone through all my boring time when I was struggling with my research. I really enjoy studying and discussing with all of them.

Last but not the least, I owe my deepest gratitude to my family, especially my parents, for their continuous supports over these years. This thesis is

dedicated to them.

Contents

Summary

This thesis consists of three chapters. Chapter 1 and Chapter 2 are completed by my own. Chapter 3 is a joint work with Professor Alberto Salvo.

Chapter 1 examines mechanism of resource allocation within households using the Brazil Family Expenditure Survey. Using monetary income share[1] as a measure of bargaining power, I evaluate its impacts on children's clothing expenditure share and female clothing expenditure share by estimating a Quadratic Almost Ideal Demand System. I test the income pooling hypothesis and examine whether the data supports the unitary setting or the collective setting. To deal with potential endogeneity issue, I control unobserved heterogeneity by mimicking Olley and Pakes (1996) control function approach. The results confirm with literature that using income share and expenditure share to examine resource allocation mechanism within households may lead to biased conclusions. Further, this work complements literature by suggesting such bias favors the collective setting and the approach adopted in this paper could mitigate such bias.

Chapter 2 compares married households and cohabiting households by developing and estimating a dynamic household collective model with limited

[1]Specifically, monetary income share of women.

commitment. I examine differences between union dissolution rates of married households and cohabiting households. Three differences are considered between married and cohabiting households: different matching qualities, different outside options[2], and different demographic characteristics. I quantify to what extent the different union dissolution rates between married and cohabiting households can be attributed to each factor by using Mexican Family Life Survey. Cohabitation in Mexico is more common compared with many other countries. Given emerging trend of cohabitation, conclusions drawn from Mexico could be relevant for countries where cohabitation is still not common, but is expected to boom in next few decades.

Chapter 3[3] examines four discontinuities in the ethanol content in blended gasoline fuel, mandated by Brazil's central government over the period 2010 to 2013, to test the joint hypotheses that (1) atmospheric ozone production in the Sao Paulo metropolitan area is limited by the volume and reactivity of hydrocarbons (”hydrocarbon limited”), and (2) increased ethanol use in the gasoline-ethanol vehicle fleet leads to higher ozone concentrations in urban Sao Paulo's ambient air. We adopt a regression discontinuity design (RDD) and flexibly test each discontinuity separately. Our finding that ozone levels

[2] In general, this is referred as utility from living alone.

[3] This paper has been submitted and is forthcoming in *Journal of the Association of Environmental and Resource Economists*

actually increased with ethanol penetration on each of the four occasions is consistent with a recent empirical study that used different identifying variation, and contrasts with a modeling study of Sao Paulo's atmosphere that predicted significant ozone abatement from hypothetical ethanol use. We find no significant relationship between ethanol versus gasoline use and PM2.5 levels. Current tailpipe emissions standards prescribe the exclusion of the mass of unburned ethanol that is emitted. Our result suggests that this standard should be reviewed. Following decades of hydrocarbon emissions control, urban areas in the US and elsewhere that are currently hydrocarbon-limited may see ozone levels rise if and when they adopt mid-level ethanol gasoline blends, whether to meet the Renewable Fuel Standard or Intended Nationally Determined Contributions to abating fossil-fuel emissions agreed upon at COP-21 in Paris.

List of Figures

List of Tables

Chapter 1

Resource Allocation Within Households: Evidence from the Brazil Family Expenditure Survey

1.1 Introduction

Resource allocation within households has been examined extensively in literature. The unitary setting considers a household as that which functions like an individual and makes optimal decisions. For example, it aggregates individual preferences into some kind of social preference function (Samuelson, 1956), or considers that there is a dictator of sorts within the family (Becker, 1974, 1991). The non-unitary approach assumes each member in a household has her own preference. These individuals bargain over consumption allocation according to various mechanisms, which could be categorized into two general frameworks: cooperative models and non-cooperative models. On the one hand, the cooperative models assume that outcomes within a household are Pareto efficient. There are models based on the axiomatic theory of bargaining with symmetric information such as the Nash Bargaining Model, which maximizes the product of individual surplus (Manser and Brown, 1980; McElroy and Horney, 1981). Individual surplus is defined as the difference between utility within the household and outside options, such as utility from divorce or living alone. There is also the collective model which does not impose any assumption on specific bargaining form (Chiappori 1988, 1992). Considering that household members interacting with each other closely, the efficiency assumption seems intuitive. On the other hand, the non-cooperative models use the Cournot-Nash equilibrium concept, and the outcome within a household could be inefficient. Outside options for individuals are defined as outcomes of non-cooperative games. When household members cannot reach consensus, they may act non-cooperatively, rather than getting divorced. The existence of abuse within a household may also support such models.

A growing body of literature rejects the unitary model for intrahouse resource allocation. Males and females tend to have different preferences and different ways of allocating resources. One important factor for the non-unitary setting is the bargaining power of individuals. With higher bargaining power, individuals can get more resources within a household. However, with the unitary model, there is no such concept as the household makes decisions as one individual. Researchers use different indicators to measure the bargaining power of individuals within a household and examine how this affects resource allocation within households. In general, women are found to allocate more resources to health related products, nutrition, and children as they gain more bargaining power, whereas men tend to allocate more on private consumption products like tobacco and alcohol (Lundberg, Pollak and Wales, 1997; Li and Wu 2011). Such a fact is important for policy makers when they consider using policy intervention to improve children's welfare or equality within households.

Control over economic resources has been considered an important and reliable indicator for bargaining power within a household. One idea is to examine whether the source of income within a household; that is, who earns the income, affects resource allocation. In the unitary setting, the source of income does not matter, whereas in the non-unitary setting, the source of income matters.[1] Labor income and non-labor income have been used extensively to represent the bargaining power of individuals. However, they may not be exogenous to the outcomes under investigation. Labor supply and consumption could be jointly determined. Some researchers use non-labor income such as

[1]In the unitary setting, a household makes decision as one individual. It pools all the income together. Hence, the source of income does not matter for resource allocation. However, in the non-unitary setting, individuals within a household are allowed to have different preferences. Individuals with higher income can allocate more resources according to their own preferences.

assets brought into the house when married (Thomas, Contreras, and Frankenburg 2002). Other researchers use policy interventions. Lundberg, Pollak and Wales (1997) examined the effects of changing the recipient of welfare benefits from the man to the woman in the United Kingdom, which is determined outside the domain of the household. Some other measurements have also been suggested to indicate bargaining power, such as marriage market sex ratio, and divorce law (Chiappori, Fortin, and Lacroix 2002), as well as gender of children (Li and Wu 2011). Marriage market sex ratio and divorce law may affect individuals' outside options; for example, the probability of getting married again after divorce or financial benefits one can get from divorce. Hence, it can affect individuals' bargaining power within a household. Regarding gender of children, certain societies may have strong preferences for a specific gender of children. By having that gender of children, female bargaining power within household could strengthen.

In this paper, I adopt the commonly applied assumption in this literature that resource allocation within household is efficient. Hence, I focus on the unitary model and the cooperative model. For the cooperative model, I use the collective model which is more general in the sense of not imposing any restrictions on bargaining form. The unitary model can be considered as a special case for the collective model. I use monetary income share of household members as an indicator for bargaining power. I deal with the potential endogeneity issue by employing control function approach. The framework is based on Bourguignon, Browning and Chiappori (2009), and Browning and Chiappori (1998). Under the unitary setting, monetary income share should not affect resource allocation within households. Under the collective setting, monetary income share affects the resource allocation. My main objectives are to test whether the data supports the unitary setting or the collective setting within a household using income share and expenditure share, and provide new empirical evidence. There are two reasons for considering monetary income share of individuals as an indicator for bargaining power. First, this indicator is reliable and widely applicable. It is directly correlated with control over economic resources within households. Some indicators may only work for specific societies or regions; for example, gender of children. Secondly, this indicator is widely and publicly available, for example, from census data for many countries, as well as the consumption data. However, as suggested by literature, this indicator may not be exogenous to the outcome variables under investigation, which in this case, is consumption. One contribution of this work deals with the issue of endogeneity, and here, I refer it as individual unobserved characteristics

. I employ control function approach proposed by Olley and Pakes (1996) to control for unobserved characteristics, which helps solve the endogeneity issue between labor supply and consumption. To the best of my knowledge, this is the first work to adopt such an approach in studying the mechanism of household resource allocation.

The data is based on the Brazil Family Expenditure Survey. By using the Quadratic Almost Ideal Demand System, I examine how female income share affects budget allocation toward clothing within households. Without controlling unobserved individual characteristics, as female income share increases from 0 to 1, children's clothing expenditure share increases by 8 percentage point. Female clothing expenditure share increases by 8.6 percentage point for households with one child, and 8.4 percentage point for households without any children[2]. All the results are significant. The unitary setting is rejected. However, after controlling for unobserved individual characteristics, children's clothing expenditure share only increases by 5.4 percentage point. Female clothing expenditure share increases by 3.6 percentage point for households with one child, and increases by 7.9 percentage point for households with no children. Moreover, all the results are insignificant and the unitary setting cannot be rejected. Without controlling for unobserved individual characteristics,

[2]To be more specific, without children means with no children living in households. Couples could have children, but they may live outside of households. This applies to remaining sections of this paper.

the effects of female income share on clothing expenditures tend to be overestimated, which favors the collective setting. Hence, I concur with literature that using income share and expenditure share to explore the mechanism of intrahouse resource allocation may result in misleading conclusions. Further, my findings complement existing literature by showing that the results bias towards the collective setting. I find that the control function approach can mitigate such bias. I also explore the effects by considering rich households and poor households, rural households and urban households, and households in both north and south Brazil. The effects vary among the different groups, and suggest that different types of households may require different indicators and outcome variables for examining the mechanism of resource allocation within households.

The remaining sections of the paper are as follows: Section 2 reviews relevant literature, including that relating to the collective setting and demand system. Section 3 describes the data. Section 4 discusses potential endogeneity issues. Section 5 explains the model, and Section 6 discusses the estimation strategy. Section 7 lists the results, and Section 8 concludes.

1.2 Literature review

The unitary model assumes a household makes decisions like an individual. According to Samuelson (1956), one may aggregate individual preferences within a household into the following household utility function:

$$\tilde{U}(Q, q^a, q^b) = H(U^a(Q, q^a, q^b), U^b(Q, q^a, q^b)) \tag{1.1}$$

where Q is public good. q is private consumption. a and b indicate different individuals. H is strictly increasing in individual utilities, U^a and U^b. The demand functions derived from such household utility function satisfy Slutsky condition, adding-up, and homogeneity. Moreover, it preserves income pooling, which means the distribution of income within a household does not affect resource allocation within the household. Another possible mechanism is due to Rotten Kid Theorem (Becker, 1974, 1991). If there is an altruistic adult in the household, who has much more income compared with other household members and can make monetary transfer, then the household members will internalize decisions and behave like one individual.

The non-unitary model considers the fact that household consists of several individuals. Decisions are made through the interactions among each member. It could be a solution of bargaining game; for example, the Nash bargaining models (McElroy and Horney, 1980; Manser and Brown, 1980). There exists a threat point for each member, which represents the maximum utility one can get from some default outcome, such as divorce. Some other researchers consider that the default outcome need not be divorce. The threat point could be the a non-cooperative outcome within a household (Lundberg and Pollak, 1993). The outcome may also satisfy some form of efficiency. The collective model, first proposed by Chiappori (1988), assumes that outcomes within a household are Pareto efficient, without any assumption about the specific bargaining form. Each individual has her bargaining power, which is important in determining the final outcome. There are multiple sources that can affect individual power such as relative income, age, divorce law, etc. The power of individuals is correlated with the aforementioned threat point. The household maximizes a weighted sum of each individual's utility. The weighting functions can be considered as indicators of the bargaining power of each member. Another stream of literature assumes that the outcome may also be a solution to non-cooperative games, which could be inefficient. One such model is the separate sphere approach of Lundberg and Pollak (1993).

Empirically, a growing body of evidence rejects the unitary setting. The wide-spread rejection of income pooling in empirical literature has been influential in weakening the case for the unitary model (Lundberg and Pollak, 2008). Using data of Canadian couples with no children, Browning, Bourguignon, Chiappori, and Lechnen (1994) find that final allocations of expenditures on each partner depend significantly on their relative incomes and ages, and on the level of lifetime wealth. They examine the expenditure sharing rule within a household. By exploiting the policy change in United Kingdom that changes the children's benefit recipient from male to female, Lundberg, Pollak, and Wales (1997) find that there is a substantial increase in spending on women's and children's clothing, relative to men's clothing, which coincides with this income redistribution. Attanasio and Lechene (2002) examine the PROGRESA program in rural Mexico. This program transfer money to poor households in rural Mexico. The implementation of this program was delayed in a number of randomly chosen villages, and hence, can be viewed as a randomized experiment. They find that as the wife's income share increases, there is greater expenditure on food and children's clothing, and less expenditure on services and alcohol. Using the Indonesian Family Life Survey, Thomas, Contreras and Frankenberg (2002) examine the effect of distribution of power within family on children's health. They use individual assets brought into a household when married in order to measure bargaining power. Their results suggest that in Java and Sumatra households, where mothers are more powerful, daughters have fewer episodes of illness than sons. Based on the prevalent son preference in China, Li and Wu (2011) construct a measure of bargaining power using gender of children. They find that women with first-born sons have a greater role in household decision making, and are less likely to be underweight. Using South Korea data, Lee (2007) examines the relationship between pocket money and income share within a household. This paper finds bargaining power is stable within a household and is insensitive to income share once household fixed effect is controlled. However, across households, bargaining power varies with income share. The author attributes these findings to unobserved bargaining power when the households are formed, which contributes to the correlation between bargaining power and income share across households. Once such unobserved bargaining power is controlled, within a household, household members can commit to the pre-specified resource allocation plan, and the income pooling hypothesis cannot be rejected. Hence, the unobserved bargaining power may play an more important role in determining resource allocation. These works focus on consumption. One typical difficulty is tracking consumption identity from household level consumption data. Excludable and assignable[3] goods can be used to solve this issue. For example, we may assume male clothes are consumed solely by males and female clothes are solely consumed by females. For tobacco and alcohol, men are more likely

[3] There is no strict difference between excludable goods and assignable goods. For example, male clothes and female clothes are excludable goods as they are consumed exclusively by corresponding parties. Clothes is assignable good. We can assign male clothes to men and female clothes to women.

to consume these than women.

There are also works examining labor supply. Using Panel Study of Income Dynamics data for the 1988 cohort, Chiappori and Lacroix (2002) find that state level sex ratio and divorce laws that favor women affect labor supply behavior and decision process in the direction predicted by theory, and have sizable effects. For example, passage of a divorce law that favors women will induce more money transfer from husbands to wives. Blundell et al. (2007) estimate a collective labor supply model using data on married couples without children, derived from the UK Family Expenditure Survey during the years 1978 to 2001. They consider potential wage for males, and find that increases in the potential wage of husband increases the wife's labor supply.

In general, empirical counterparts reject the unitary setting in many different countries. They also support the prediction of collective model in the sense that individuals with greater bargaining power can get more of the resources they prefer. There is also another stream of literature examining whether the resource allocation within households is efficient. The evidences are mixed (Bobonis, 2009, Udry, 1996) and beyond the scope of this work.

In this work, I focus on the collective setting and assume the resource allocation within a household is efficient based on the usual assumption adopted by the collective model. I examine the intrahousehold resource allocation mechanism by examining the relationship between income share and clothing expenditure share. I adopt control function approach to deal with the endogeneity issue.

1.3 Data

The data used in this work is derived from the Brazil Family Expenditure Survey. The survey period is from May 2008 to May 2009 covering 27 states in Brazil. In my study, I only consider households that consist of a couple with no children or a couple with one child younger than 15 years old in order to control potential confounding effects from children. Households with multi-families are also excluded. There are two reasons to restrict to households with a child younger than 15. Firstly and foremost, the survey only keeps separate track of clothing expenditures for children younger than 15. For children older than 15, their clothing expenditures are combined with their parents' clothing expenditures depending on their sex. Second, as children grow older, they may develop certain bargaining power within households (Dauphin, Lahga, Fortin and Lacroix, 2011). This may complicate identification. Therefore, I examine households with one child and households with no children living in households separately. These two types of households could be different in terms of resource allocation.

The main outcome variable is clothing expenditure share[4] for each member in the household. I divide the child's clothing expenditure by total clothing expenditure within the household to generate the child's clothing expenditure share. I divide female clothing expenditure by adult clothing expenditure[5] to generate the female clothes expenditure share. Zero expenditure for each subcategory is allowed. However, if the total clothing expenditure is zero, I have to drop the data. It is unclear how to define budget share in this way. The final sample contains 10368 households. Of this total, 5133 households have one child living in the household, and 5235 households have no children living in the household. The summary statistics are in Table A.1. Of the sample, 23.9 percent of households live in rural area. More than half (56.7 percent) of households are in the north part of Brazil. Compared with the south part of Brazil, the north part is less developed. 92.8 percent of husbands supply labor and 67.0 percent of wives supply labor. With regard to ownership, 35.5 percent of households own cars, and 60.2 percent of households own houses. As suggested in literature, car ownerships and house ownerships are important factors affecting household consumption. To control wealth level, I adopt

[4]One concern is clothes may show nature of public good. For example, husband may care about clothes of wife. This would not be an issue as we can model this case using caring preference. In collective setup, with some mild assumptions, we can transform a problem of caring preference to a problem of egoistic preference.

[5]Adult clothes expenditure is summation of male clothing expenditure and female clothing expenditure.

a point-based system (Eizenberg, A and Salvo, A., 2015). The points are calculated based on four factors: bathrooms and sewage system, household head education, asset holding, and servants. Detailed information regarding the points calculation is in Appendix C.

1.4 Model

1.4.1 Decision making process

Before I discuss the collective setup, I will introduce the decision-making process for my model. The decision process is presented in Figure B.1. A and B denote household members. L is the amount of labor supply. τ is unobserved characteristics, which is the main issue I intend to deal with. m is income. q denotes private consumption, and Q is public consumption. Each Household decides amount of labor supply and consumption. These two decisions are made sequentially, as each household enters the labor supply decision period first, and then consumption decision period. First, in the labor supply decision period, which is represented by the hollow points in the figure, the household decides amount of labor supply of each member, which depends on unobserved individual characteristics τ and observed individual characteristics[6]. Then, the household moves to the consumption decision period. Along the line denoting

[6]The observed characteristics are omitted in the figure, as well as in the following analysis for clarity. They can be incorporated in the model easily.

the consumption decision period in the Figure, there are several solid nodes which indicate the time when a household makes a consumption decision. The income depends on the labor supply decision of respective household members, but not on unobserved characteristics conditional on labor supply. Given each individual's income, the household makes consumption decisions that depends on income as well as unobserved characteristics. After several consumption decisions, the household may reconsider labor supply and enter a labor supply decision period once again, which is represented by the hollow point on the right in the Figure.

Several points should be further clarified. First, in real life, a household makes consumption decisions more frequently than labor supply decisions. The aforementioned model illustrate this fact by considering many consumption decision periods following one labor supply decision period. In my analysis, I focus on the consumption decision period and consider income as fixed when a household makes a consumption decision. I will not be examining labor supply decisions. Second, over time, I do allow for simultaneity between labor supply decisions and consumption decisions. Both decisions are affected by unobserved characteristic τ. A natural question to address is the economic meaning of τ, which should directly affect labor supply decisions and consumptions decision. One example we may consider is the relative preference of consumption over preference of leisure. If an individual has a relatively strong preference of consumption over leisure, in general, this individual may tend to supply more labor and also consume more goods. In the aforementioned sequential game, this means that this individual will supply more labor in the labor supply decision period. Also, given the income, this individual will also tend to consume more goods[7]. This story also explains why failing to consider unobserved characteristics may lead to misleading conclusions later, when I consider bargaining power within the household. Thirdly, the unobserved characteristics could be correlated between husband and wife within a household. For example, the unobserved characteristics could be positively correlated due to positive assorting in the initial matching stage. This assumption is not necessary for the derivation of the later section. However, it may affect the degree of bias if we do not consider these unobserved characteristics. Considering the previous example, in the extreme case when such unobserved characteristics are exactly the same for husband and wife within one household, there will be no bias if the unobserved characteristics are not considered. In real life, a certain degree of correlation of unobserved characteristics between husband and wife should be expected.

[7]This point will be further clarified in model part

1.4.2 Collective setting

The collective setting allows household members to have different utility functions. The assumption for such model is that resource allocation within a household is Pareto Efficient. The model can be transformed into the following program:

$$
\begin{aligned}
&max_{q^A,q^B,Q}\mu(\mathbf{p},m)u^A(q^A,q^B,Q)+[1-\mu(\mathbf{p},m)]u^B(q^A,q^B,Q)\\
&\qquad st:\mathbf{p}(q^A+q^b+Q)=m
\end{aligned}
\tag{1.2}
$$

Here, the household maximizes household utility, which is a weighted sum of individuals' utility functions[8], u^A and u^B, subject to budget constraints. q^A and q^B represent private consumption of each individual. $\mathbf{p}$ represents price and m is income. Q represents consumption in public good. $\mu(\mathbf{p},m)$ and $(1-\mu(\mathbf{p},m))$ represent weights of each individual. A natural explanation for μ is bargaining power of each individual. Following Browning and Chiappori (1998), I assume there exists a differentiable, zero-homogeneous function $\mu(\mathbf{p},m)$ such that for any $(\mathbf{p},m)$, the vectors $(\mathbf{q}^A,\mathbf{q}^B,Q)$ are solutions to the this program. The assumption postulates that the decision process always has a unique, well-defined outcome, or there exists a demand function.

Demographic information can be added into this formula. Some factors could enter into the utility function $u(\bullet)$ directly, such as ownership of house and car, age, and education. Another groups of factors, distributing factors,

[8]Usually, we refer to utility of head and spouse within households.

enter into the weight function $\mu(\bullet)$. They affect consumption decisions within a household by altering the bargaining power of household members. They do not enter into the individual utility function.

In this work, I incorporate unobserved characteristics of individuals into the model and examine how this may affect the conclusion of the resource allocation mechanism within household. The modified program is as follow:

$$\begin{aligned} & max_{\{q^A,q^B,Q\}} \mu(m^A(\tau^A,\tau^B),m^B(\tau^A,\tau^B)),\mathbf{p})u^A(q^A,q^B,Q,\tau^A)+ \\ & \qquad [1-\mu(m^A(\tau^A,\tau^B),m^B(\tau^A,\tau^B)),\mathbf{p})]u^B(q^A,q^B,Q,\tau^B) \qquad (1.3) \\ & \qquad st: \mathbf{p}(q^A+q^B+Q)=m^A(\tau^A,\tau^B)+m^B(\tau^A,\tau^B) \end{aligned}$$

This program demonstrates decisions made by a household in consumption period as in Figure B.1. At this period, the income[9] of each member is provided, and household only makes consumption decisions. If panel data is available and τ is fixed over time, omitting τ may not cause bias if identification comes from variation of consumption and income within the household. However, in my case, when only cross sectional data is available, omitting τ may bias the results and favor the collective model.

The unobserved characteristic τ can affect final consumption through three potential channels. The first channel is utility function, which I refer to as di-

[9]One may notice the income here is $m(\tau^A,\tau^B)$, and in previous decision making process, the income is $m(L)$. The full form should be $m(L(\tau^A,\tau^B))$. Unobserved characteristics affect income only through labor supply. Labor supply decision depends on unobserved characteristics, but is determined previously, not in consumption period. Hence, to avoided the potential confusion that treating labor supply decision as endogeneous in this period, I omit L here.

rect effect. The second is income effect. In the labor supply decision period, τ affects the labor supply decision, and hence, income available in the consumption period. The third channel is bargaining power effect. Since τ affects the individual labor supply decision, it will also affect the income distribution within a household in consumption period. If income distribution affects individuals bargaining power and hence the weighting function μ, τ will also affect the consumption decision. I refer to the second and third channels as indirect effect.

In the empirical counterpart, I will examine the relationship between consumption budget share and income share across households. Without controlling unobserved characteristic τ, the second effect should not cause a problem as we can control income directly. However, the first and third channels will cause problems. The first channel indicates that unobserved characteristic are correlated with the dependent variable, consumption. The third channel indicates that unobserved characteristic are correlated with independent variable, income share. To be consistent with my empirical part, I made the following change to the above program:

$$
\begin{aligned}
max_{\{q^A,q^B,Q\}}\mu(s(\tau^A,\tau^B))u^A(q^A,Q,\tau^A)+ \\
[1-\mu(s(\tau^A,\tau^B))]u^B(q^B,Q,\tau^B) \\
st:\mathbf{p}(q^A+q^B+Q)=\bar{m} \\
where:\bar{m}=m^A(\tau^A,\tau^B)+m^B(\tau^A,\tau^B) \\
s(\tau^A,\tau^B)=\frac{m^A(\tau^A,\tau^B)}{\bar{m}}
\end{aligned} \tag{1.4}
$$

s is the income share for A, which I will use to examine bargaining power in the empirical analysis. As I focus on the consumption decision period, I fix the total income within households, which can simplify the analysis below. Income effect can be controlled directly in the empirical analysis. Individuals are assumed to have egoistic preferences and the utility function is concave in terms of private and public goods. There are three assumptions for unobserved characteristics τ:

Assumption 1: $\frac{\partial u^A}{\partial \tau^A}>0$, $\frac{\partial u^B}{\partial \tau^B}>0$

Assumption 2: $\frac{\partial^2 u^A}{\partial q^A \partial \tau^A}>0$, $\frac{\partial^2 u^B}{\partial q^B \partial \tau^B}>0$, $\frac{\partial^2 u^A}{\partial Q \partial \tau^A}>0$, $\frac{\partial^2 u^B}{\partial Q \partial \tau^B}>0$

Assumption 3: $\frac{\partial m^A}{\partial \tau^A}>0, \frac{\partial m^A}{\partial \tau^B}=0$, $\frac{\partial m^B}{\partial \tau^B}>0$, $\frac{\partial m^B}{\partial \tau^A}=0$.

Assumption 1 states that fixing a consumption bundle, individual utility is an increasing function of unobserved characteristic τ. Hence, an individual tends to consume more as τ increases. Assumption 2 states that individual marginal utility is an increasing function of unobserved characteristic

τ. Assumption 3 states that an individual supplies more labor and earns more income as unobserved characteristic increases. However, a spouse's unobserved characteristics have no direct effect on the person's own labor supply. A spouse's unobserved characteristics can still have an indirect effect on individual decision of labor supply; for example, through consumption. Next, I will examine potential bias, the calculation for which is in Appendix D.

As total income is fixed at consumption stage, given Assumption 2, τ^B can be expressed as a function of τ^A[10]. We denote it as $\tau^B = \phi(\tau^A)$. Denote q^{A*}, q^{B*}, and Q^* as the optimal consumption bundles for aforementioned program. We can get equations[11] below. The key part to distinguish the unitary setting and the collective setting is $\frac{d\mu}{ds}$. $\frac{d\mu}{ds} > 0$ rejects the unitary model and favors the collective model as income share affects bargaining power. If $\frac{d\mu}{ds} = 0$, the unitary model cannot be rejected.

$$\frac{\mathbf{d}q^{A*}}{\mathbf{d}s} = \frac{\partial q^{A*}}{\partial \mu}\frac{\mathbf{d}\mu}{\mathbf{d}s} + \frac{\partial q^{A*}/\partial \tau^A}{\mathbf{d}s/\mathbf{d}\tau^A} \tag{1.5}$$

$$\frac{\mathbf{d}q^{B*}}{\mathbf{d}s} = \frac{\partial q^{B*}}{\partial \mu}\frac{\mathbf{d}\mu}{\mathbf{d}s} + \frac{\partial q^{B*}/\partial \tau^A}{\mathbf{d}s/\mathbf{d}\tau^A} \tag{1.6}$$

$$\frac{\mathbf{d}Q^{*}}{\mathbf{d}s} = \frac{\partial Q^{*}}{\partial \mu}\frac{\mathbf{d}\mu}{\mathbf{d}s} + \frac{\partial Q^{*}/\partial \tau^A}{\mathbf{d}s/\mathbf{d}\tau^A} \tag{1.7}$$

In Equation 1.5, 1.6, and 1.7, the left-hand side is what we can observe from the data. These could be the evidence suggesting the unitary model or

[10] Refer to Appendix D.

[11] Any thing that can be controlled empirically is omitted for the equations below, for example, the income effect.

the collective model. Focusing on Equation 1.5, if $\frac{\mathbf{d}q^{A*}}{\mathbf{d}s}$ is positive, this means individual A consumes more as her income share increases, which suggests the collective model. However, the term $\frac{d\mu}{ds}$ on the right-hand side of the equation is direct evidence suggesting the collective model or the unitary model. If $\frac{d\mu}{ds}$ is positive, this means income share affects individual weight, or bargaining power within a household. Hence, the unitary setting is rejected. The second term on the right-hand side of the equation determines the bias if $\frac{\mathbf{d}q^{A*}}{\mathbf{d}s}$ is used. The exact direction of bias depends on the magnitude of the first order derivative of utility function, which is nontrivial to examine. Moreover, imposing assumptions on the relative magnitude of first order derivative of utility function has little empirical implication. Hence, I consider the following simplified cases.

Begin by assuming that there are no public goods in the household, and concentrate on private consumption of each individual. Focusing on Equation 1.5 first, $\frac{\partial q^{A*}}{\partial \mu}$ is positive. Basically, a higher weight favors the individual and enables her to consume more. Hence, $\frac{\partial q^{A*}}{\partial \mu}\frac{d\mu}{ds}$ is either positive or 0, which indicates the collective setting or the unitary setting respectively. The remaining term on the right-hand side $\frac{\partial q^{A*}/\partial \tau^{A}}{\mathbf{d}s/\mathbf{d}\tau^{A}}$ is positive, which is the bias if we do not control unobserved characteristics of individuals. Even if $\frac{d\mu}{ds}$ is zero, which means the weight or the bargaining power of individual does not depend on income share, we still observe that an individual consumes more as her income share increases. The concept is illustrated by the following: Considering two households 1 and 2, they are identical except that individual A_1 has a higher unobserved characteristic[12] in household 1 while individual A_2 has a lower unobserved characteristic in household 2. In the labor supply decision period, A_1 will supply more labor compared with A_2, and hence obtains a higher income share within the household. Conditional on income, in the consumption period, A_1 will also consume more private goods compared with A_2. Comparing these two households, we can find a positive relationship between individuals' income share and consumption within the household. However, the effect does not go through bargaining within the household. Hence, we may falsely conclude the unitary setting as the collective setting without control for unobserved characteristics of individual.

Second, for the case in which public goods are considered, I discuss from these two perspectives: stage budgeting and relative preference for public goods. The stage budgeting assumption states that consumers can allocate budgets in stages. In this work, I adopt two-stage budgeting assumption. To be more specific, consumers allocate resources to public goods first, then they allocate resources among private goods. For the second stage, when consumer

[12] One may consider high unobserved characteristic as high valuation for consumption or lower valuation for leisure.

allocates private goods, the analysis of bias for private consumption is the same as no public goods case, as public goods expenditure is fixed. Hence, the bargaining is within the context of private consumption. For the direction of bias of public goods, which is the second term $\frac{\partial Q^*/\partial \tau^A}{\mathbf{d}s/\mathbf{d}\tau^A}$ in Equation 1.7, this depends on individual relative preference for public goods. Consider the extreme case in which individual A cares about public goods and individual B derives no utility from public goods, the bias term $\frac{\partial Q^*/\partial \tau^A}{\mathbf{d}s/\mathbf{d}\tau^A}$ will be positive. On the other hand, if individual A does not care about public goods and individual B cares about public goods, then the bias term $\frac{\partial Q^*/\partial \tau^A}{\mathbf{d}s/\mathbf{d}\tau^A}$ will be negative. The direction of bias depends on empirical context. In my empirical analysis, I consider the effect of female income share on children's clothing expenditure. As suggested by literature, women tend to care more about children than do men. Therefore, intuitively, the bias term should be positive. This means without considering unobserved characteristics, the results favor collective models.

In general, from these simplified cases, without controlling unobserved characteristics, we get biased results when we use income share and expenditure share to examine the mechanism of household resource allocation. With cross-sectional data, the bias may lead us to accept the collective model when the true model is the unitary model. For public consumption, the bias is less clear and depends on the empirical context. In the empirical segment, I will impose stage budgeting assumptions to make the situation parallel to these simplified cases.

1.4.3 Dynamic issue

In my model, I assume static efficiency while making no assumptions as to dynamic efficiency. Basically, within the consumption decision period, I assume that allocation within a household is efficient. However, overtime, I do not impose any restrictions on dynamic efficiency. There are at least two explanations in literature regarding the existence of dynamic inefficiency. First, household members may over supply labor to boost their bargaining power in the consumption period (Basu, 2006). Second, a household may have limited ability to stick to a pre-specified plan, and instead makes decisions based on current state (Mazzocco, 2007). The first explanation does not contradict nor support my model. Essentially, my model does not need restrictions on the efficiency of labor supply. The second explanation can support my model. One problem people may have is that a household may make labor supply decision and consumption decision simultaneously, not sequentially. This does not cause problem for my model. It is possible that the household makes both decisions jointly. However, when they make consumption decisions later, they cannot commit to the pre-specified plan, and make decisions based on current state. Essentially, these two streams of literature do not invalidate my model.

1.5 Empirical Model

1.5.1 Demand system

As mentioned previously, there exists a demand function that can represent the aforementioned model. I use the Quadratic Almost Idea Demand System proposed by Banks, Blundell and Lewbel (1997). This demand system is flexible and gives an arbitrary first-order approximation to any demand system

$$\begin{aligned} w_j = \delta z_j + \alpha(x_j, \tau_j) + \Gamma \mathbf{p}_j + \beta(x_j, \tau_j)(lnE_j - lna(x_j, \tau_j, \mathbf{p}_j)) \\ + \lambda \frac{(lnE_j - lna(x_j, \tau_j, \mathbf{p}_j))^2}{b(x_j, \tau_j, \mathbf{p}_j)} \end{aligned} \tag{1.8}$$

$$lna(x_j, \tau_j, \mathbf{p}_j) = \alpha_0 + \alpha(x_j, \tau_j)'\mathbf{p}_j + \frac{1}{2}\mathbf{p}_j'\Gamma\mathbf{p}_j \tag{1.9}$$

$$b(x_j, \tau_j, \mathbf{p}_j) = exp(\beta(x_j, \tau_j)'\mathbf{p}_j) \tag{1.10}$$

$$\alpha(x_j, \tau_j) = \alpha^0 + \alpha^1 x_{j1} + \alpha^2 x_{j2} + \alpha^3 x_{j3} + ...\alpha^s x_{js} + \alpha^\tau \tau_j \tag{1.11}$$

$$\beta(x_j, \tau_j) = \beta^0 + \beta^1 x_{j1} + \beta^2 x_{j2} + \beta^3 x_{j3} + ...\beta^s x_{js} + \beta^\tau \tau_j \tag{1.12}$$

w_j is the vector of budget share of goods for household j. z_j is the vector of distributing factors. x_j is the vector of observed demographic information, $(x_{j1}, x_{j2}, ..., x_{js})$, that enters the utility function directly. Additionally, τ is unobserved characteristics that enter the utility function. $\mathbf{p}_j$ is the vector of log price faced by household j. E_j is the total expenditure for goods in consideration.

I incorporate x and τ nonlinearly into the system, and incorporate z linearly into the system. There are two reasons for doing this. First, there is one proportional results from the theoretical model of the collective setting (Browning and Chiappori, 1998):

$$\frac{\partial w_{j1}/\partial z_1}{\partial w_{j1}/\partial z_2} = \frac{\partial w_{j2}/\partial z_1}{\partial w_{j2}/\partial z_2} \tag{1.13}$$

The simple way to incorporate or test such a restriction is to model z linearly. The second reason is for identification of z. Basically, I need linearity in z and nonlinearity in τ to separate the effect of z on consumption from the effect of z on τ through control function.

1.5.2 Control function

Following Olley and Pakes (1996), I model consumer unobserved heterogeneity using a scalar in the empirical segment. Based on the collective model in previous section, consumption can be written as:

$$(q^{A*}, q^{B*}) = h(x, z, m, p, \tau^A, \tau^B) \tag{1.14}$$

Conditional on x, z, m, p, q^{A*} is increasing in τ^A,[13] and decreasing in τ^B; while q^{B*} is increasing in τ^B, and decreasing in τ^A. This is the same as the monotonicity assumption of Olley and Pakes. Hence, I can invert this equation

[13] Refer to Part 7 in Appendix D.

and obtain the following:

$$(\tau^A, \tau^B) = \hat{h}(x, z, m, p, q^{A*}, q^{B*}) \tag{1.15}$$

Then, I plug this equation above into the right-hand side of the demand system by replacing τ with $\hat{h}$:

$$\alpha(x_j) = \alpha^0 + \alpha^1 x_{j1} + \alpha^2 x_{j2} + ...\alpha^s x_{js} + \alpha^\tau \hat{h}(x_j, z_j, m_j, p_j, \tilde{w_j}) \tag{1.16}$$

$$\beta(x_j) = \beta^0 + \beta^1 x_{j1} + \beta^2 x_{j2} + ...\beta^s x_{js} + \beta^h \hat{h}(x_j, z_j, m_j, p_j, \tilde{w_j}) \tag{1.17}$$

The basic idea here is to use one group of goods to invert for unobserved characteristics, and use another group of goods to examine the variable in which I am interested, or the test income pooling hypothesis. I use clothing expenditure share to test the income pooling hypothesis, and use the personal service expenditure share[14] to control for unobserved individual characteristics.

One assumption I need to validate using the demand system and control function approach is weak separability. This assumption states that preferences for products of one group are independent of product-specific consumption of products from other groups. To be more specific, conditioning on total expenditure of clothes and total expenditure of personal service, the way that a household allocates expenditures among clothes is independent of the way that the household allocates expenditures among personal service. This as-

[14]For example haircut, massage, etc.

sumption reduces the dimension of analysis, which means I can concentrate on examining clothing expenditure without considering other expenditures. Hence, it is valid to use expenditure of personal service to control for unobserved characteristics and plug into the demand system of clothes. Otherwise, I should consider expenditures in all cases as dependent variables.

The weak separability assumption has been commonly used for previous literature, especially when the researchers examine a specific category of consumption, such as food, or clothes. Under the framework of current work, one potential issue could involve the possibility of collective trade off and bargaining at a higher level. For example, it is possible that a female may require less food for herself and argue for more clothes. Such an argument can be applied to all categories of consumption. To fully address this issue, I must consider all the consumption categories in the regression which would be costly. Moreover, it is challenging to identify such collective trade off behavior across different categories of goods from general preference, as much data are not available; for example, personal consumption of food. Finally, in current work, this would not be a major threat to my results as long as the collective bargaining at a higher level is not systematically correlated with income share. For example, it is hard to rationalize a female with higher income share systematically arguing for more female clothes by consuming less food, or vice versa.

1.5.3 Estimation

I use the iterated linear least square estimator proposed by Blundell and Robin (1999) to estimate the system. I parameterize $\hat{h}$ linearly. As mentioned, I consider two groups of consumption in the empirical segment, clothes and personal service. I use personal service to control for unobserved characteristics, and use clothing expenditure to examine the mechanism of resource allocation. To be more specific, I examine whether female income share within a household affects expenditure on female clothing, and children's clothing. The parameter of interests is δ. I examine the model using two steps: I consider the expenditure share of children of all clothes purchased within a household, and I consider the expenditure share of the female's clothing over all adult clothing purchased within a household[15]. Basically, I examine the expenditure on public goods first, and then examine the bargaining for private consumption[16]. For the controlled variables in the demand system, distributing factors include female income share [17], age and education difference of husband and wife, and head sex dummy. Demographic information entering the utility function include monthly household total income, female age and education, children's

[15]Adult clothes is the summation of male clothes and female clothes.

[16]This process can be validated by two stage budgeting assumption.

[17]I use female income share in demand function as distributing factor and use log income ratio in control function as distributing factor. The main reason is to reduce standard error and improve precision of estimation.

age and sex if applicable, labor supply decision, car ownership, house ownership, household points, clothing prices, rural fixed effect, monthly fixed effect and state fixed effect.

1.6 Results

Table A.2 presents the results of female income share on clothing expenditure for households with no children and households with one child. Column (1) is for households with a child and column (2) is for households with no children. Results without control for unobserved characteristics are presented in Panel A, and results with control for unobserved characteristics are in Panel B.

If unobserved individual characteristics are not controlled, as female income share increases from 0 to 1, budget share for children's clothes increases by 8 percentage point. Female clothing expenditure share increases by 8.6 percentage point for households with one child, and by 8.4 percentage point for households without children. As a female earns relatively more income within her household, she has higher bargaining power. Hence, she can allocate more resources to her own. Moreover, similar to previous literature, the female also allocates more resources to her child as she gains more bargaining power within the household. These results support the collective setting.

After controlling for individual unobserved characteristics, all the effects become smaller. Children's clothing expenditure share increases by 5.4 percentage point, as female income share increases from 0 to 1. Female clothing expenditure share increases by 3.6 percentage point for households with one child, and increase by 7.9 percentage point for households with no child. The results are consistent with the prediction based on the model; without controlling unobserved characteristics, we may overestimate the effects of income share on expenditure share. The effect of income share on female clothing expenditure share changes much smaller for households with no children, about one-tenth of the change for the households with one child, after control for unobserved individual characteristics. This means that omitting individual unobserved characteristics has a smaller effect on the bias for households with no children. One explanation could be, if I consider unobserved characteristics as a relative preference of consumption over leisure, household members without children have more free time available. If marginal utility from consumption of leisure decreases, with enough time available, individuals will supply labor eventually even if they have low unobserved characteristics, or strong preferences for leisure. For households with a child, free time is more limited as members have to spend time taking care of the child. For those who have a low unobserved characteristics, or a high preference for leisure, they will supply less labor, compared with counterpart members[18] in a household with no children. In general, we may expect the bias to be higher for households whose members' free time is more limited.

Considering the significant level, all the results become insignificant after unobserved individual characteristics are controlled. This means we cannot reject the unitary setting. However, if we compare standard errors in two cases, the standard errors become larger when unobserved individual characteristics are controlled. This could be due multicollinearity as both demand system and control function contain indicators constructed by income as distributor factors. Hence, the inability to reject the unitary model could be due to lack of power here.[19] Future studies should attempt to address this issue. One potential way to solve this issue is to find two groups of goods that depend on different distributing factors.

Clothing expenditure makes up a small part of total household expenditure. For the sample I use, the mean share of clothing expenditure of monetary monthly income is about 13.8 percent within households. For poor households, the share is 14.4 percent, which is higher than the 11.8 percent spent by rich households. We may expect such bargaining could be more likely to

[18]We can think two exactly same individuals, but one in household with a child, and one in household without children.

[19]The standard errors for the following results also show similar pattern.

occur in poor households. Resources in these households are more limited and their members have more incentive to bargain over individual consumption for clothes. Table A.3 presents rich-level specific effects[20]. In general, the effects are higher for poor households compared with rich households. No matter whether unobserved individual characteristics are controlled or not, all the effects are insignificant for the rich households. Hence, the unitary setting cannot be rejected for the rich households. However, we should be cautious when interpreting the results. Based on clothing expenditure, individuals in rich households seem not to bargain; however, they may still bargain about expenditures of other more expensive goods.

In the empirical segment, I only consider monetary income and use it to construct income share to measure the bargaining power of individuals. One concern is that a household may still have non-monetary income; for example, crops or home production, that is not marketable. Income share constructed using monetary income may not be a good measurement of control over economic resources. Hence, I consider urban and rural specific effect. The assumption is that in urban areas, households are less likely to have non-monetary income, or that the monetary income is the major component of household income. Hence, income share using monetary income is a good in-

[20]Based on household points constructed, top 20% are considered as rich households and remaining are considered as poor households.

dicator of control over economic resources. For rural areas, households are more likely to have nonmonetary income, or nonmonetary income is the larger component of household income. The monetary income share will be a noisy measurement for bargaining power. Table A.4 presents the urban and rural specific effects

In the rural area, the effects are very small. After unobserved individual characteristics are controlled, the effect becomes negative for female clothing budget share in the household with a child. First, monetary income is a noisy measurement of bargaining power for the rural household. Hence, the sign may not make any economic sense. Second, it is possible that there is a negative relationship between monetary income and nonmonetary income within a household; for example, each member focuses on one part. Nonmonetary income could be a major source of income in rural areas. A negative relationship could be possible in this case. For urban households, the effects are much bigger as compared with rural households; this is as expected. Hence, when using control over economic resources as an indicator for bargaining power within households, researchers should be careful about the source of income for households with different demographic characteristics.

In Brazil, the southern part of the country is more developed than northern part. Based on the sample, in the northern region, 33.6 percent of households are from rural areas, whereas 20 percent of households are from rural areas in southern part. Hence, it is interesting to examine whether there are any regional differences. I categorize the sample into south and north based on the state. The south consists of the south region, southeast region, and central region of Brazil. The north consists of the north region and northeast region. The results are presented in Table A.5.

Comparing north and south, the effects on female clothing expenditure share are similar for both households with one child and households with no children. However, the effect is different for children's clothing expenditure share. The effect is much greater for households in the south region. One explanation could be that monetary income is not a good indicator for northern region. For example, the northern region consists of more rural areas. However, it is unclear how this explains the different patterns of effect on children's clothing expenditure share and female clothing expenditure share. Another explanation could be that children in the north region may contribute to household income earlier than children in the south. Since these areas are less developed, it is possible that children may also have to supply labor[21]. In this case, children themselves may have bargaining power and ability to allocate resources to themselves. Considering this as a possibility, there would

[21] It is possible that the income is not enough and children have to supply labor. It is also possible the return of education is low.

be less incentive for the female to further allocate resources to her children.

In all the cases above, after unobserved individual characteristics are controlled, the effects of female income share become smaller and tend to be insignificant. This is important for policy makers to consider when they make monetary transfers to households. Distribution of benefits to different recipients could lead to different final resource allocations, depending on the recipient; however, the difference may not be huge. Moreover, the effects are heterogeneous across households. Usually, such policies target poor households. For households in the north region, if policy makers intend to improve children's welfare, it seems that such a policy is indifferent between different recipients based on my results. However, such a policy may also change other factors; for example, children are more likely to withdraw from the labor market if the female receives the money. This is outside of the scope of this work.

One concern is how robust the results derived from this paper may be, due to the restrictive sample. In this work, I focus on households with either one child or households with no children. One may question whether conclusions derived from this paper can be extended to households with multiple children. In the theoretical model part, the discussion is insensitive to number of children. Hence, potential bias due to ignorance of unobserved characteristics could also exist in households with multiple children. The assertion that we may extend to the whole population is to be careful with such bias. Moreover, the method adopted in this work can be extended to households with multiple children.

1.7 Conclusion

In this work, I examine the mechanism of household resource allocation. The primary framework I adopt is the collective setting, and the unitary setting can be considered as a special case. This study assumes outcome within a household is Pareto efficient. I modify the model by incorporating unobserved individual characteristics. I show how omitting these variables could cause biased results, and favor the collective setting over the unitary setting if income share is considered as the indicator for measuring bargaining power. The bias of effect on private consumption is clear. However, for public goods, the bias is unclear and depends on the empirical settings.

I then parameterize the model using the Quadratic Almost Ideal Demand system and test the prediction using Brazil data. I adopt control function approach by Olley and Pakes to control for unobserved individual characteristics. In general, after unobserved characteristics are controlled, effects of female income share on clothing expenditure for both children and females decrease and become insignificant. The empirical results do support the theoretical prediction. Previous literature suggests that regarding money transfer policies, by transferring money to the female, more will be allocated to children, compared with giving money to the male. My results do not completely reject such a suggestion, however; the difference may not be that huge.

Based on my empirical content, I further study heterogeneous effects by considering rich households versus poor households, urban households versus rural households, and households in the north part of Brazil versus households in the south part of Brazil. Depending on the demographic characteristics of households, different measures for bargaining power and outcome variables should be selected for examining the resource allocation mechanism within households. This is also important for government officials when considering the implementation of cash transfer policies. Depending on the goal of the government, households with different demographic characteristics may react differently. Policy makers should recognize their main target and consider the effects of such policies accordingly. Finally, this paper focuses on Brazil; however, the method adopted in this work can be applied to a wide range of data.

Chapter 2

Stability of Married Households and Cohabiting Households

2.1 Introduction

Cohabitation has gained popularity as a potential substitute for marriage over the past few decades. This trend has emerged across different continents. The meaning of cohabitation without formal marriage depends on the social context. Some societies have considered cohabitation a common practice for a long time. For example, many Latin American countries have long histories of socially accepted consensual unions, which may substitute for formal unions in some groups (De Vos, 1999; Parrado and Tienda 1997). In the European Union, cohabitation is common in many countries. For example, premarital cohabitation in Sweden is nearly universal. In many northern European countries, cohabitation has progressed further into the direction of becoming a replacement for marriage (Pollak and Lundberg, 2013). In some other places, cohabitation has become acceptable, emerging quickly in recent decades. In the US, cohabitation has increased from roughly 500,000 couples in 1970 to more than 7.7 million couples in 2010. This growth is evident across racial and ethnic, education, and age groups (Lofquist, Lugaila, O'Connell, Feliz, 2012).

Many factors contribute to this emerging trend. Individuals can get benefits from cohabitation similar to marriage, such as economies of scale, risk sharing, consumption of public good, and overcoming credit constrain, but these benefits may be to varying degrees. Moreover, in many countries, cohabiting individuals are required to take less responsibility toward their partner and the union, both from the perspective of law and social convention. There are fewer legal concerns when cohabiting couples want to separate compared with separation of married couples. But still, this varies country by country. For example, in Canada, common-law partners, referring to couples that live together without an actual ceremony or legal documents, have the same fundamental rights as married couples after two years of cohabitation. In most Asian countries, there is no such law protecting cohabiting couples. Some people may also use cohabitation as a trial period for marriage. The social norm has changed over time making people more tolerant of premarital sex and living together without getting married. The stigmas associated with nonmarital sex, cohabitation, nonmarital fertility, and divorce have declined dramatically (Thornton and Young-DeMarco, 2001). With more options available, people also delay marriage nowadays. Some people value the freedom of their leisure time. Others want to pursue their own career. Women in particular have more opportunities in the labor market than in the past. Getting married is more likely to hamper a woman's career development compared with the career development of a man. At the same time, individuals may also desire certain levels of intimacy. Cohabitation could be a better choice compared with marriage. Finally, people may also face financial constraints when they want to get married, such as expenditure related to wedding or dowry. In such cases, people might choose cohabitation as an alternative.

The standard economic model assumes only two living arrangement for households, getting married or living alone (Becker, 1981, 1991, Weiss, 1997). This would have been a reasonable assumption in the past. However, with cohabitation becoming more popular, such an assumption may not be adequate and might over-simplify the analysis of households' behaviors and decision making. One natural question to address would be whether it is necessary to model cohabiting unions and married unions separately. Scholars have different views about behaviors between married households and cohabiting households. Some consider that behaviors between households of both types of living arrangement; that is, marriage and cohabitation, should converge over time. The second demographic transition theory posits that cohabitation should become more normative and look increasingly like marriage over time. Childbearing should become more common for cohabiting households and cohabiting families should become more stable over time (Kiernan 2000; van de Kaa 1987). With the trend that social norm defining married partners' behavior becomes weaker and social norm defining cohabitation becomes stronger, the deinstitutionalization hypothesis also predicts that the distinction between marriage and cohabitation should fade over time (Cherlin, 2004). One opposing view suggests the increasing privilege of marriage, due to people self-selecting into marriage and the social status, which may make behaviors of cohabiting households and married households diverge from each other, such as with regard to stability of households (Cherlin 2009; Furstenberg 1996). Many countries are still in a transitional stage. Depending on social norms and cultural convention, it is also possible that different countries may transit along different directions. For most countries, cohabitation is still not viewed as a perfect substitute for marriage. Existing research shows that cohabiting couples are less stable than married couples. Cohabiting households are less likely to invest in long term public goods, such as a house or children (Pollak and Lundberg, 2013). These facts themselves could be correlated. For example, the instability of cohabiting households could be the reason for lower investment in long term public goods. The lower investment in public goods within households may further contribute to the instability of these cohabiting households. Hence, behaviors for married households and cohabiting households are different along certain dimensions and potentially will remain so for the next few decades.

Past studies suggest that economically disadvantaged groups, such as low education and low income individuals, are more likely to cohabit. Upon union dissolution, these people could be prone to financial problems. With children bearing activities becoming more common within cohabiting unions, children from cohabiting households may suffer more from union dissolution compared with children from married households. This could be due to the economic disadvantage of cohabiting households. Many countries do not have laws protecting children from cohabiting households in case of union dissolution.

In this paper, I examine the stability of cohabiting households and married households, including reasons for the difference between the divorce rate for married households and separation rate for cohabiting households. One complication facing this study is that married households and cohabiting households are different along many dimensions. I categorize these differences into three groups: unobserved matching qualities, outside options, and demographic characteristics. Unobserved matching qualities measure how well the male and female partners are matched with each other. These are observed by the male partner and female partner, but not observed by the econometrician. On average, married couples and cohabiting couples may have different unobserved matching qualities for two reasons: different initial selection process and different in-house investment after unions are formed[1]. Outside options are referred as individuals' utility if unions are dissolved. They measure how well individuals can live on their own[2]. Given one married household and one cohabiting household with exactly the same demographic characteristics at individual level, the outside options for individuals from these two living arrangement could be different. In most countries, cohabiting couples and married couples face different regulations when unions are dissolved. Such regulation could affect individual outside options; for example, through asset allocation.

I examine to what extent these three factors, unobserved matching qualities, outside options, and demographic characteristics, can explain the differences in stability between married households and cohabiting households. I model these three factors together by considering a dynamic household collec-

[1]This work cannot separate the effects from these two factors.

[2]This applies for cooperative model, which is considered in this paper. For non-cooperative model, outside options could mean outcome of non-cooperative game, not necessarily living alone.

tive model with limited commitment. I quantify the effect of different factors using the Mexican Family Life Survey.

The remaining of the paper is organized as follows: Section 2 presents a literature review, and focuses on reasons explaining differences between married households and cohabiting households. Section 3 shows data and reduced form analysis. Section 4 describes the model. Section 5 presents results and section 6 concludes.

2.2 Literature review

Cohabitation can afford similar benefits as marriage, such as joint consumption and production. However, behaviors between cohabiting households and married households are different in certain dimensions. One important difference is in the stability of households. Cohabiting unions are found to be less stable compared with married unions. Today's cohabiting unions are less likely to culminate in marriage and more likely to end in separation. Even if cohabiting couples eventually move to marriage, premarital cohabitation is associated with a higher subsequent risk of divorce, as suggested by many studies (Balakrishnan et al. 1987; Bennett, Blanc, and Bloom 1988; Hoem and Hoem 1992; Bracher et al. 1993; Lillard, Brien, and Waite 1995; Weiss and Willis 1997; Brien, Lillard, and Stern 2006). Compared with married households, cohabiting households are less likely to invest in long term public good, such as a house or children (Pollak and Lundberg, 2013). There are several possible explanations for the aforementioned differences in households' behaviors.

Firstly, married households and cohabiting households face different costs for union dissolution. Such cost could generate differences in households' behaviors, both directly and indirectly through intertemporal commitment. The legal process and cost for dissolution of union greatly distinguish marriage from cohabitation. It is more costly for married households to dissolve their union compared with cohabiting households. Moreover, social convention may impose an additional psychological cost on married couples if they get divorced. Cohabiting households may just consider their arrangement a temporary or informal union. Individuals may care about a failed marriage, but are less likely to care about a failed cohabitation. These factors could enhance the stability of married households directly compared with that of cohabiting households. Due to the lower cost of dissolution of a union and informal property from cohabitation, the intertemporal commitment generated within a cohabiting union is limited. Moreover, within marriage, divorce law usually favors the more vulnerable side, in most cases women, by allocating more assets or requiring alimony transfer. This further enhances the intertemporal commitment within marriage. To understand the persistence of marriage once cohabitation is recognized as an alternative, the economic approach emphasizes the potential returns to intertemporal commitment (Pollak and Lundberg, 2013). Intertemporal commitment can provide credit that facilitates long-term investment, both for tangible assets (e.g, house) and intangible assets (e.g, education), and risk pooling (Weiss, 1997). Devices ensuring intertemporal commitment are important, especially for women, as women have been specialized to childbearing and other domestic activities, leaving them vulnerable to abandonment and other adversities (Becker, 1991). Regulated by law and social convention, marriage itself could serve as a contract or device that guarantees the intertemporal commitment within households (Matouschek, Rasul 2008). As such, the intertemporal commitment would be strengthened and more cooperation could be expected within married households. Married households are more likely to invest in public goods within their households. These additional benefits further ensure the stability of married households.

Secondly, couples that select themselves into marriage or cohabitation could be different in both observed and unobserved ways. In the US, cohabitation is strongly decreasing in education (Pollak and Lundberg, 2013). Serial cohabitation is much more prevalent among economically disadvantaged men and women (Lichter and Qian, 2008). As suggested by Pollak and Lundberg (2013), there is an education and income gradient between the decisions of marriage and cohabitation. One explanation is the competing force of decreasing return from labor specialization and increasing return from investment in children. Market substitutes for housework, such as cooking and cleaning, are more easily accessible and affordable today. Additionally, with more opportunities for women in the labor market, traditional gains from marriage such as joint consumption, production and labor specialization decrease over time. For investment in children, researchers find that parents with higher education and income spend more time and money on children (Ramey and Ramey, 2010). The elasticity of expenditures on many children items is higher than one. Moreover, recent works also document the complementarity between early human capital stock and the productivity of later investment in children (Aizer and Cunha 2012; Cunha and Heckman, 2007; Todd and Wolpin, 2003, 2007). Thus, these results suggest the return from investment in children will be much higher for households with higher education and income. Marriage would be a more optimum environment for raising children; for example, due to the aforementioned intertemporal commitment. Thus, for high education and income households, the increasing return from investment in children outweights the decreasing return from labor specialization, making marriage a more attractive option. The contrary is the case for the low education and income households, which makes cohabitation a better option. It is also possible that individuals prefer to be matched with high education and income partners. Such chance would be small for low income and education individuals. Hence, some of them may form temporary cohabiting unions with other low education and income partners, while searching for a better match.

Cohabiting households and married households could also differ in unobserved ways. People who select to cohabit may not be ready for a stable relationship. This may also make the households less likely to invest in long term assets. Such unobserved heterogeneity due to self-selection could also be one explanation for the fact that marriages preceded by cohabitation are less stable (Lillard, Brien, and Waite 1995). Another reason for couples to choose cohabitation could be that they don't have enough information about their partners. Hence, there will be greater uncertainty with cohabiting. Moreover, it is also possible that the couple realize their matching qualities are low, and hence they choose cohabitation. Similar to any long-term investment, married households are also more likely to invest in pursuits of matching qualities, such as more family bonding activities, which improves the potential for matching qualities compared with cohabiting households. Susan et al. (2015) compared the relationship quality of cohabitors and marrieds. They found that the relationship between union type and relationship quality is bifurcated, with direct marrieds reporting the highest relationship quality and cohabitors without marriage plans reporting the lowest marital quality. Marrieds who premaritally cohabited and cohabitors with plans to marry fell in the middle.

Thirdly, outside options could be different for married couples and cohabiting couples. We can also consider this one unobservable factor that makes individuals self-select into marriage or cohabitation. Outside options are usually referred as individuals' utility when couple get divorced and separated, or utility from being single. Such options are considered the threshold of in-household utility for couples under cooperative models, which could reflect individuals bargaining power and resource allocation in households. If current resource allocation scheme within a household makes one particular individual's utility lower than the threshold, that individual can initiate negotiation to increase her resource share within the household. If no such option can be found, divorce or separation will be triggered. This framework applies to cohabiting households as separation can be triggered unilaterally. However, for married households, this outcome depends on whether the law states unilateral divorce or mutual consent divorce. Under the mutual consent regime, the union can survive as long as one of partners prefers staying together over living alone. Under the unilateral regime, resource allocation within households needs to guarantee that both members prefer staying together over living alone. In this work, following the law in Mexico, a unilateral regime for married households is adopted. The utility from living alone could depend on individuals education, income, assets and other demographic characteristics. It could also depend on unobservable factors such as an individual's attractiveness, confidence, or ability to network. One common factor considered in the literature affecting outside options is income. One possibility for instability of cohabiting households could be that couples have better outside options. It is more difficult for them to reach consensus in terms of resource allocation that keeps both of them preferring to stay within the union rather than separating. However, this explanation seems to contradict the literature that contends that economically disadvantaged groups are more likely to choose cohabitation, but have worse outside options. Another possibility could be that these individuals may generate more gains from forming unions. And cohabitation is more easily formed than marriage.

In this work, I examine three main factors: unobserved matching qualities, outside options, and demographic characteristics. The effect of demographic characteristics on outside options can be modelled directly. Unobserved matching qualities are modelled as a random process. For connection between demographic characteristics and unobserved matching qualities, I focus on education based on literature and allow that households with different education levels have different random processes for unobserved matching qualities. Given demographic characteristics, unobserved matching qualities and outside options are independent in the model. Policy intervention; for example, imposing divorce law on cohabiting households, can change outside options in the short run, but not unobserved matching qualities. However, such a policy intervention may change unobserved matching qualities in the long term[3]. Through the counterfactuals, I fix demographic characteristics. I change the unobserved matching qualities, and asset division rule which changes outside options of individuals. I then examine how the separation rate for cohabiting households changes.

A dynamic household collective model with limited commitment can incorporate these factors and would be a suitable framework. This model was proposed and tested by Mazzocco (2007). The collective model, formalized by Chiappori (1988,1992) is a special case of the Nash bargaining model (Manser and Brown, 1980; McElroy and Horney, 1981), whereas households maximize a weighted sum of utility across each member. The weight reflects individuals' bargaining power, which could depend on individuals' outside options. The original works for the collective model are static. The dynamic model is necessary here as one important channel to explain the results is through

[3]With similar regulations imposed on cohabitation and marriage, people may consider these two living arrangements similar. The observed and unobserved demographics of these two groups may converge over time. However, this issue is not examined in this work.

intertemporal commitment. Married households and cohabiting households may invest differently within households; for example, buying a house, purchasing cars, given different probability of separation. And such investments may affect surplus generated within households. Moreover, I want to examine movement between living arrangements, among marriage, cohabitation, and divorce or separation. Full commitment assumes that households consider all the possible uncertainties in the future and design a consumption plan at the beginning of union formation. This is rejected by Mazzocco (2007) and limited commitment is favored. The limited commitment model assumes that in states where individuals utility within households are lower than outside options, renegotiation within households will be triggered. The model adopted in this paper is similar to that of Voena (2015), where she examines how divorce law change in the US affects married household members' behavior. Another similar work is by Gemici and Laufer (2011), who consider cohabitation using a similar model with the National Longitudinal Survey for the US. However, they focus on the efficiency issue within households while this paper focuses on union stability.

2.3 Data and Reduced Form Analysis

The data set used in the empirical analysis is from the Mexican Family Life Survey. This panel data set consists of three waves: the years 2002 to 2003, 2005 to 2006, and 2009 to 2012[4]. The initial sample from 2002 consists of 8400 households in 150 urban and rural communities. In this work, I focus on households that were already formed with one male partner and one female partner in 2002. One of the partners has to be the head of household. Both the male partner and female partner have to be living in the household in the year 2002. The male partner should be aged at least 25 years old and younger than 55 years old. In total, there are 3352 eligible households in year 2002, with 2848 married households and 504 cohabiting households. Cohabitation is not uncommon in Mexico, even back to the 1990s. In the 2000 Mexico census and 2010 Mexico census, there were 23% and 38% cohabiting unions among all unions, respectively. Cohabitation has become more common over time.

The summary statistics for the households in 2002 are illustrated in Table E.1. Male partners and female partners in cohabiting households are on average 2 years younger than partners among married couples. In this year, the average number of years cohabiting in the sample is 13.5 years. The average

[4]For the last panel, the majority of interviews were completed in 2010. Due to tracking issues, some households were only reached until year 2012.

number of years married in the same sample is 18 years, 4.5 years longer than cohabiting households. The age difference may explain part of the difference in number of years of union formation. The average age for men when forming a union is 22 for married households and 24.5 for cohabiting households. In my empirical analysis, I assume all unions are formed when male partners are 25 years old. Cohabiting households have fewer children than do married households. In this work, I do not model decisions on having children, and consider the number of children as given. For education, I consider two categories: having finished high school or above, and not having finished high school. Similar to other literature, the average education level is higher for married households than for cohabiting households. In married couple, 24% of males and 16% of females finished high school, while 18% of males and 10.5% of females finished high school in cohabiting households. Married households have about 70% more asset than do cohabiting households. There could be several explanations for such differences in assets. First, the difference could partially be explained by the difference in education level. Second, couples may require a certain level of asset possession in order to get married. Third, married households may invest or save more as mentioned previously due to intertemporal commitment. The income difference is smaller than the difference in assets. Males in married households earn 10% more than in cohabiting households, and the difference is insignificant. However, females earn 54% more in cohabiting households than in married households. And conditional on the female who earns income, females in cohabiting households still earn 35% more than the females in married households. Facing a higher separation rate, women in cohabiting union may have more incentive to maintain an income source, as they will need money to survive if they separate from their male partner. In general, the data is consistent with literature that states that cohabiting households are more likely to be in economically disadvantaged groups, with lower education levels and lower asset values.

Of the 2848 married households, 2546 (89.40%) households stayed married throughout the period of study (three waves of panel). 101 (3.55%) households changed to cohabitation. 19 households in the 101 households changed to married again[5]. The remaining 201[6] (7.05%) households got divorced. For the 504 cohabiting households, 224 (44.44%) households stayed cohabiting over the period of study. Another 157 (31.15%) households got married later, and 58(11.51%) households got married, but changed to cohabiting again. Finally, 65 (12.90%) households got separated. In general, married couples are less likely to get divorced (separated) compared with cohabiting couples. Ta-

[5]This behavior seems odd. There could be some errors or typos in the data. However, the sample size is small and the results are not sensitive if we assume these households stayed married for entire survey period.

[6]For households with changing partners, they are considered as divorce or separation in previous period.

ble E.2 shows results examining separation and marriage rates for cohabiting households based on 2002 demographics using multinomial logistic regression. Households are considered separated if they separated in either 2005 or 2009 panels. Households are considered married if they are married for both 2005 and 2009, or cohabiting in 2005 and married in 2009. All the other households are considered status quo, which is the base outcome in the regression. Female income and male asset share[7] are positively correlated with separation rate. These two variables could be indicators of outside options. As mentioned previously, individuals facing higher separation rates may deliberately maintain higher outside options, and higher outside options make individuals more likely to separate. Total assets and number of children are negatively associated with separation rate. These two factors can be considered in-house investments. Households can gain from such investments and hence are relatively more stable. Female education is positively correlated with probability of cohabiting households getting married. One interpretation could be that a high education female can manage family bonding better, which may help households move to marriage. Table E.3 shows results examining divorce rates for married households. Households are considered divorced if they got divorced in either 2005 or 2009 panels[8]. Female income is positively correlated

[7]Discussions on generating male asset share are presented later.

[8]Households changing to cohabitation are considered in the same category as households

with divorce rate. Male education and female age are negatively correlated with divorce rate. Based on this analysis, education and income affect movement among households' living arrangement. In the empirical part, I consider households with different education levels separately. The effect of income can be examined in the model, as it can enter budget constraint. However, effect of education is hard to model directly. One way could be to model income process differently for different education levels[9], which may not be enough to capture the differences from education. Another way is to assume that matching qualities differ between different education levels. Hence, I conduct the empirical analysis separately based on education level of household members.

Literature usually considers three categories for education: combining college graduate and some college education, high school graduate, and not finishing high school. Due to sample size concerns, I only consider two categories for education: finishing high school and above, and not finishing high school and below. For households, I consider two categories for education match: households with both partners not finishing high school, and households with at least one finishing high school. I refer to the first group as low education households and the second group as high education households. In 2002, there

staying married. The goal for this regression is to examine divorce behaviour. Moreover, moving to cohabitation still requires certain level of cooperation between male partners and female partners.

[9]I do model income process differently based on education level.

are 2409 low education households, with 2022 (83.94%) of them being married households and 387 (16.06%) cohabiting households. There are 943 high education households, with 826 (87.59%) married households and 117 (12.41%) cohabiting households in 2002. I assume high education households and low education households have different matching qualities, for both cohabitation and marriage. A high education male or female may care more about qualities with their partners and spend more effort searching for potential partners before the union is formed. Moreover, education itself could serve as a signal for matching qualities, which could positively affect matching qualities as well as reduce variance of matching qualities. Hence, at the beginning, matching qualities for high education households could be better than that for low education households, both in terms of higher mean and smaller variance. After the union is formed, high education women may enlist more effort to improve matching qualities within households. This is to say, even if the matching qualities start purely randomly for high education households and low education households, high education households may invest more in family bonding. Overtime, matching qualities would become better for high education households. The results shown in Table E.2 and E.3 may support these arguments, as education of both the male and female have a positive effect on household stability.

To model union dissolution, I need to consider how households allocate assets and potential money transfer between partners. For married households in Mexico, the asset allocation depends on the judge. Females may expect to get 50% to 70% of assets in a divorce, depending on the economic status of each individual. For cohabiting households, I assume they take the share of assets belonging to them. The specific question I use to infer the asset share from the survey is in the section of survey entitled Household Asset. The question asks: If the household had to sell or make use of the (...), how many people would make the decision of doing it? Then, respondents have to indicate the exact members who would make such a decision, including: themselves, their partners, children, parents, parents in law, brothers and sisters, and brothers in law and sisters in law. If certain members can make such a decision, I assume they own part of the asset. I only consider the male partner and female partner for ownership of the asset. For a specific asset, the share of ownership for male partner is 1 if he can make the decision on selling and using while the female partner cannot. The share of ownership is 1/2 if both male partner and female partner can make the decision. A similar idea applies to female partner. The categories of assets considered in the questionnaire include; for example, houses, domestic applications, trucks, and so on[10]. Then, I calculate individual owned share of total assets by considering the value and share of ownership for male and female partners for each asset. One might worry about the reliability of this indicator as I do not observe the exact share of each asset. On the one hand, the indicator generated here could be considered as the true asset share plus noise. Two assumptions to validate using this indicator are, firstly, there is a positive correlation between true asset share and the indicator I use; secondly, the noise is purely random and not correlated with other factors. On the other hand, results in Table E.2 and Table E.3 may partially justify the validity of this indicator. Male asset share affects the separation rate for cohabiting households, but not divorce rate for married households. Another type of regulation to consider is money transfer between partners; for example, alimony. I do not model this in the current work for two reasons. Firstly, I do not observe such money transfer in the data. Secondly, considering the case of money transfer from men to women, such a transfer can be stopped if the woman gets remarried. This complicates tracking money transfer in the model. Hence, in this work, I assume married women get 70% of assets during a divorce, to compensate for the fact that I do not consider money transfer.

[10]There are in total 7 categories: 1. House occupied by this household including land; 2. Other house, construction, real estate, plot, agricultural ground or land, 3. Bicycle; 4. Motorcycle, automobile, truck, or any other motorized vehicle; 5. Electronic appliances (radio, TV set, VCR, Computer, etc.) 6. Washing machine and dryer, stove, refrigerator, and furniture; 7. Electronic and domestic appliances (blender, iron, microwave, etc.))

In the following section, I examine income and time allocation for households and how these evolve over time. These are related to the moments I use to estimate the model in a later section. The main objective is to demonstrate that married households and cohabiting households are similar in these aspects and that it is reasonable to consider these two living arrangement using one model.

Households are categorized into five groups: 1. Households that stayed married for three waves of interview; 2. Households that stayed cohabiting for three waves of interview; 3. Households that changed from married to cohabiting; 4. Households that changed from cohabiting to married, but not the other way around[11]; 5. Households that divorced or separated.

Figure F.1 shows the proportion of male partners having incomes among all households. Figure F.2 shows the proportion of female partners having incomes among all households. In 2002, more than 90 percent of male partners had incomes. The proportion is similar in 2005, except in the case of married households that changed to cohabiting. There, the proportion dropped to 85 percent. In 2009, there is a drop in the proportion of male partners that have incomes. One potential explanation could be the negative economic shock

[11]For example, a cohabiting household in 2002, that changed to married in 2005 and changed back to cohabiting in 2009, is not in this category. This household is in previous category.

in Mexico in 2009. The GDP growth rate was 1.4% in 2008, and dropped to -4.7% in 2009. The proportion of male partners that have incomes is the highest for households cohabiting for all three waves of interviews. The lines appear parallel between households remaining cohabiting and married. In 2002, 28% to 37% of female partners have incomes. The proportions are stable for households that stay married and cohabiting in 2005. If there is any change, the proportions drop. This could suggest more labor specialization over time. For divorced or separated households, the proportion increases to 45%, as now, female partners are living on their own. They have much stronger incentive to find an income source. In 2009, the proportion of divorced women that have income still increases. For households that changed from being married to cohabiting, and from cohabiting to being married, they seem follow the previous trend with slightly decreasing proportions. For households that stayed married and cohabiting over three panels, the proportion of female partners that have incomes increases in 2009, even despite the negative economic shock. One explanation could be the add-work effect. As male partners lose jobs or their income source, the female partners will go out to find jobs or an income source to support their family. As mentioned previously, the proportion for male partners having income drops for all five groups of households. But potential for the add-work effect only exists for households staying cohabiting or married, which may suggest more cooperation within these two types of households. In 2002 and 2005, households staying married or cohabiting for the entire study period have the lowest proportion of female partners having incomes. This may also suggest the higher degree of labor specialization within these two types of households. The lines appear parallel between households staying cohabiting and staying married. To further examine the time allocation within households, Figure F.3 shows the average number of working hours per week for men, and Figure F.5 shows the average number of working hours per week for women. Figure F.4 shows average housework hours per week for men and Figure F.6 shows average housework hours per week for women. The graphs suggest labor specialization in all types of households, with men spending more time working and women spending more time on housework. Notable is that the women's hours of housework dropped for all types of households over time. One explanation could be a better production asset within households. In the household production function I consider in later section, I model household production function depending on production asset value to account for decreasing female housework hours.

To summarize, based on summary statistics and reduced form analysis, cohabiting households are less stable than married households. Female income is negatively correlated with household stability. Male asset share decreases stability of cohabiting households, but not that of married households. All types of households show labor specialization, with the households continuing staying married and cohabiting showing this to the greatest extent. Education also interacts with living arrangement. Compared with low education households, a greater proportion of high education households choose marriage. Moreover, over time, high education cohabiting households are more likely to result in marriage, with female education playing a significant role. Male education decreases the divorce rate within married households. In next section, I develop a model to further examine these effects.

2.4 Model

The model I consider is a household collective model with limited commitment. Male partner and female partner jointly decide savings, labor supply and household production. This model focuses on households that are already formed. Only married households and cohabiting households are examined. I do not consider the matching process between male and female partners.

The household lives from period 1 to period T. Male partner and female partner work from period 1 to period T-R, and retire from period T-R+1 to period T. In each period, each household chooses from the following three living arrangements: being married, cohabiting, or getting divorced or separated.[12].

[12]In this work, divorce means that husband and wife divorce and no longer make decisions

2.4.1 Preferences

Male and female partners derive utility from leisure l_i, and composite good Q_i produced within household, for $i = H, W$[13]. If the male partner and female partner choose to stay together, either married or cohabiting, they derive a utility θ_t^j, which is observed by the couple but unobserved by econometricians, depending on their living arrangement, where $j = M, C$[14]. Thus, the flow utility for male partner and female partner in each period can be expressed as follows:

$$\begin{aligned} u_{it}^{Married} &= u(l_{it}, Q_{it}) + \theta_t^M \\ u_{it}^{Cohabit} &= u(l_{it}, Q_{it}) + \theta_t^C \\ u_{it}^{Divorce} &= u(l_{it}, Q_{it}) \end{aligned} \tag{2.1}$$

Utility from leisure and composite good are assumed to be separable. I use a CRRA utility function for composite good and linear utility function for leisure as shown in Equation 2.2. There is an additional penalty for men if they do not work. h_{Ht}^w denotes hour of work for men. $\mathbb{I}_{h_{Ht}^w=0}$ is indicator function. ϕ is a number between 0 and 1. Men lose part of the utility from leisure if they do not work. There are two potential explanations. Firstly, as men are usually

as one household. If a household changes from being married to cohabiting, this means they divorce but still make decisions as a household. If a household changes from being married to divorced, this means they no longer make decisions as a household. Similarly, if a cohabiting household changes to separated, this means male and female partners no longer make decisions as a household.

[13] H indicates male partner and W indicates female partner

[14] M indicates married and C indicates cohabiting.

considered the main income source within households, they may suffer certain psychological costs if they do not work. In such cases, they may not fully enjoy their leisure time. Secondly, if men do not work, they may spend time looking for a new job, rather than spending their time for leisure. For women, there is no such penalty if they do not work.

$$\begin{aligned} u(l_{Ht}, Q_{Ht}) &= \delta_l * l_{Ht} * (1 - \phi * \mathbb{I}_{h^w_{Ht}=0}) + \frac{Q_{Ht}^{1-\delta_Q}}{1-\delta_Q} \\ u(l_{Wt}, Q_{Wt}) &= \delta_l * l_{Wt} + \frac{Q_{Wt}^{1-\delta_Q}}{1-\delta_Q} \end{aligned} \tag{2.2}$$

The unobserved matching qualities, shown in Equation 2.3, follows random walk both for being married and cohabiting, which captures the persistency in taste within marriage and cohabitation. If a household stays in the same living arrangement, the expected matching quality for that specific living arrangement increases over time. Household members gain more from living together, and are less likely to separate or divorce. If a household moves from being married to cohabiting, or from cohabiting to being married, there will be no persistency transmitted. The process for unobserved matching qualities can be expressed as follows:

$$\begin{gathered} \theta_t^{j_t} = \mathbb{I}_{j_t=j_{t-1}} \theta_{t-1}^{j_{t-1}} + \epsilon_t^{j_t}, \quad j_t, j_{t-1} = M, C \\ \epsilon_t^{j_t} \sim N(\mu_{j_t}, \sigma_{j_t}^2) \\ \theta_0^{j_0} \sim N(\mu_{j_0}, \sigma_{j_0}^2) \end{gathered} \tag{2.3}$$

$\mathbb{I}_{j_t=j_{t-1}}$ is indicator function, which takes a value of 1 if the current period living arrangement is the same as in the previous period. ϵ_t^{jt} is random shock drawn in each period. Households draw tastes for marriage and cohabitation for every period. θ_0^{jo} is initial matching quality and follows the same distribution as random shock in each period.

2.4.2 Household production

The composite good produced within a household depends on the amount of time devoted to in-house production by male and female partners, h_i^{hw} for $i = H, W$, goods purchased from market are denoted as *input*, and household production technology which depends on household capital is denoted as cap_t.

$$Q_t = input_t^{\alpha} * (h_{Ht}^{hw} + h_{Wt}^{hw})^{1-\alpha} * log(cap_t) \tag{2.4}$$

The household production function, presented in Equation 2.4 is in Cobb-Douglas form. $log(cap_t)$ captures decreasing return from capital. In-house labors are assumed to be perfectly substitutable between men and women. In the literature, researchers usually assume labor specialisation by considering the following function form for in-house hour of work, in which $v > 1$:

$$((h_{Ht}^{hw})^v + (h_{Wt}^{hw})^v)^{1/v} \tag{2.5}$$

As it is unclear how cohabiting households and married households can be differentiated from this aspect, I do not focus on this point. The labor specialization in my model is reliant on wage. High wage individuals work outside and contribute to *input*, and low wage individuals focus more on housework. Total composite good available for the male and the female partner depends on equivalent scales e_t.

$$Q_{Ht} + Q_{Wt} = Q_t/e_t \tag{2.6}$$

The equivalent scale depends on the number of children in the household. In this model, the fertility choice is not modeled. Hence, the number of children is considered exogenous. If households choose to divorce, the amount available within households for particular individuals is[15]:

$$Q_{it} = input_{it}^{\alpha} * h_{it}^{hw1-\alpha} * log(cap_{it}), i = H, W \tag{2.7}$$

2.4.3 Income

From period 1 to period T-R, a household earns both labor income and non-labor income. From period T-R+1, a household only earns non-labor income. Non-labor income is considered exogenous in each period. Labor income depends on wage, which is arrived at according to Equation 2.8 below[16]:

$$\begin{aligned} log(wage_{ikt}) &= max\{\beta_{0ik} + \mathbb{I}_{h^w_{i,t-1}>0}\beta_{1ik}log(wage_{ik,t-1}) + \omega_{ikt}, 0\} \\ \omega_{ikt} &\sim N(0, \sigma^{\omega}_{ik}) \end{aligned} \tag{2.8}$$

$\mathbb{I}_{h^w_{i,t-1}>0}$ is an indicator function, with a value 1 if individuals supplied labor in the previous period and 0 otherwise. This part captures human capital accu-

[15] Equivalent scale applies similarly if there are children.

[16] I add 1 peso to all wages to avoid 0 when estimating the transition for wage.

mulation in the labor market. The process is gender and education specific. i denotes male partner (H) and female partner (W). k denotes education level.

2.4.4 Budget constraint

The budget constraint for married household and cohabiting household from period 1 to period T-R if they stay together is:

$$\begin{aligned} A_t + cap_t + input_t = nlincome_t + h^w_{Ht} * wage_{Hkt} + h^w_{Wt} * wage_{Wkt} \\ + R * A_{t-1} + cap_{t-1} \end{aligned} \tag{2.9}$$

Total flow of income for each period is the summation of non-labor income $nlincome_t$ and labor income. h^w_{Ht} and h^w_{Wt} are labor supplied by male partner and female partner, given wage $wage_{Hkt}$ and $wage_{Wkt}$ respectively. From period T-R+1 on, the total flow income is non-labor income only. A_t is non-production asset and cap_t is production asset. $input_t$ is expenditure on goods purchased from market.

If households move from being married to divorced, household members will split the assets according to law. There will be a separate budget for each member. The budget constrain for each member is:

$$\begin{aligned} A_{it} + cap_{it} + input_{it} = sharenli_i * nlincome_t + h^w_{it} * wage_{ikt} \\ + R * shareasset_i * A_{t-1} \\ + sharecapital_i * cap_{t-1} \end{aligned} \tag{2.10}$$

$shareasset_i * A_{t-1}$ and $sharecapital_i * cap_{t-1}$ are the amount of assets the male partner and female partner can expected to get based on law. $sharenli_i *$

$nlincome_t$ is the amount of non-labor income that belongs to the male partner and female partner. A_{it} and cap_{it} are the assets that each individual chooses in this period. The flow of income consists of labor income of each respective member and the proportion of non-labor income that belongs to respective member. If a household moves from cohabiting to separation, the budget constraint for each individual is similar as above. However, asset division depends on ownership. Finally, each household member devotes her time to leisure, housework or labor. The time constraint for each household member is:

$$l_{it} + h_{it}^{hw} + h_{it}^{w} = h_{it}^{total} \tag{2.11}$$

2.4.5 Household problem

In each period, for both married households and cohabiting households, the household maximizes the utility function which is the weighted sum of the male partner's and female partner's individual utilities. The Pareto weight is denoted by μ, which represents the bargaining power of the male partner and female partner. In this work, a unilateral divorce regime is adopted. Let s_t be the relevant state variable to be considered for each situation excluding living arrangement, in which the state variable at time t, besides living arrangement, includes: male age[17], production assets, non-production assets, asset

[17] When I solve the model, I only track male age to reduce calculation burden

share of male partner and female partner, male wage, female wage, non-labor income, equivalent scale, male education, female education, Pareto weight of male partner and female partner, and matching qualities.

2.4.5.1 Married household

In each period, a household can choose to continue being married, cohabiting, or getting divorced. Let V_t^j, $j = M, C, D$, denotes household living-arrangement specific value function. V_t denotes household value function considering optimization over the living arrangement. V_{it}^j, $i = H, W$, denotes living arrangement specific value function for the male partner and female partner. V_{it} denotes individual value function with the living arrangement being considered at the household level.

If a household chooses to continue being married, they solve the problem below:

$$
\begin{aligned}
V_t^M(s_t|M) = max_{A_t, cap_t, input_t, h_{Ht}^w, h_{Ht}^{hw}, h_{Wt}^w, h_{Wt}^{hw}, l_{Ht}, l_{Wt}} \\
\mu_t^{HM}\{u(l_{Ht}, Q_{Ht}) + \theta_t^M + \beta * E[V_{H,t+1}(s_{t+1}|s_t, M)]\} \quad (2.12) \\
+ \mu_t^{WM}\{u(l_{Wt}, Q_{Wt}) + \theta_t^M + \beta * E[V_{W,t+1}(s_{t+1}|s_t, M)]\}
\end{aligned}
$$

subject to budget constraint 2.9, time constraint 2.11 and in-house production constraint 2.4 and 2.6. β is the discounting factor. Households maximize the weighted sum of the male partner's utility and female partner's utility.

$$\begin{aligned} V_{Ht}^{M} &= u(l_{Ht}, Q_{Ht}) + \theta_t^M + \beta * E[V_{H,t+1}(s_{t+1}|s_t, M)] \\ V_{Wt}^{M} &= u(l_{Wt}, Q_{Wt}) + \theta_t^M + \beta * E[V_{W,t+1}(s_{t+1}|s_t, M)] \end{aligned} \tag{2.13}$$

If a household chooses to cohabit, they solve the problem below:

$$V_t^C(s_t|M) = max_{A_t, cap_t, input_t, h_{Ht}^w, h_{Ht}^{hw}, h_{Wt}^w, h_{Wt}^{hw}, l_{Ht}, l_{Wt}}$$

$$\mu_t^{HC}\{u(l_{Ht}, Q_{Ht}) + \theta_t^C + \alpha * E[V_{H,t+1}(s_{t+1}|s_t, C)]\} \tag{2.14}$$

$$+ \mu_t^{WC}\{u(l_{Wt}, Q_{Wt}) + \theta_t^C + \alpha * E[V_{W,t+1}(s_{t+1}|s_t, C)]\}$$

$$\begin{aligned} V_{Ht}^{C} &= u(l_{Ht}, Q_{Ht}) + \theta_t^C + \beta * E[V_{H,t+1}(s_{t+1}|s_t, C)] \\ V_{Wt}^{C} &= u(l_{Wt}, Q_{Wt}) + \theta_t^C + \beta * E[V_{W,t+1}(s_{t+1}|s_t, C)] \end{aligned} \tag{2.15}$$

subject to similar budget constraints as choosing being married.

If household members choose to divorce, each member solves the problem below:

$$V_{it}^D(s_t|M) = max_{A_{it}, cap_{it}, input_{it}, h_{it}^w, h_{it}^{hw}, l_{it}} u(l_{it}, Q_{it}) + \alpha * E[V_{i,t+1}(s_{t+1}|s_t, D)] \tag{2.16}$$

subject to individual budget constraint 2.10 and time constraint 2.11. I assume children live with the wife if divorce occurs. In such a case, the available consumption for male partner is:

$$Q_{Ht} = input_{Ht} * h_{Ht}^{hw} * log(cap_{Ht}) \tag{2.17}$$

The consumption for the female partner is

$$Q_{Wt} = input_{Wt} * h_{Wt}^{hw} * log(cap_{Wt})/e_{Wt} \tag{2.18}$$

Utilities for individuals in each case are $V^M_{Ht}(s_t|M)$, $V^M_{Wt}(s_t|M)$, $V^C_{Ht}(s_t|M)$, $V^C_{Wt}(s_t|M)$, $V^D_{Ht}(s_t|M)$, and $V^D_{Wt}(s_t|M)$. The parameters $\mu^{ij}_t = \mu^i_{t-1} + \kappa^{ij}_t$, $i = H, W$, $j = M, C$, are chosen to make sure the participation constraints below are satisfied, where κ^{ij}_t is the Lagrange multiplier associated with the respective participation constraints. Participation constraints for choosing marriage are:

$$\begin{aligned} V^M_{Ht}(s_t|M) &\geqslant V^D_{Ht}(s_t|M) \\ V^M_{Wt}(s_t|M) &\geqslant V^D_{Wt}(s_t|M) \end{aligned} \tag{2.19}$$

and participation constraints for cohabiting are:

$$\begin{aligned} V^C_{Ht}(s_t|M) &\geqslant V^D_{Ht}(s_t|M) \\ V^C_{Wt}(s_t|M) &\geqslant V^D_{Wt}(s_t|M) \end{aligned} \tag{2.20}$$

Consider households choosing being married. Given their bargaining power[18] from the previous period, households solve Equation 2.12 subject to budget constraints. Similarly, households solve the problem for choosing cohabitation and divorce. If both constraints in 2.19 can be satisfied, in that case, we say marriage is feasible and bargaining power for each member stays the same. If neither constraint in 2.19 can be satisfied, marriage is infeasible. And if only one partner is satisfied, we suppose the female partner's constraint is satisfied. In such a case, the woman prefers staying in the marriage, while the man wants to get a divorce. That household will renegotiate bargaining power, increasing the man's bargaining power to induce him to stay in the marriage. At the

[18] I may use bargaining power and weight interchangeably. They mean the same thing.

same time, the women should still prefer marriage over divorce. Empirically, we assume renegotiation continues until the man is indifferent between marriage and divorce, and at same time, the woman considers whether marriage is still preferable. If such new bargaining power can be found, marriage is still feasible. Following a similar idea, households also decide whether cohabitation is feasible. If neither marriage nor cohabitation is feasible, the household gets a divorce or separates. If only one is feasible, either marriage or cohabitation, the household chooses the feasible arrangement. If both arrangements are feasible, the household moves to the next stage, and chooses the optimal choice between these two feasible living arrangements with updated bargaining power if applicable..

If either constraint 2.21 or constraint 2.22 is satisfied, households will choose the corresponding arrangement. If neither constraint 2.21 nor constraint 2.22 is satisfied, consider the case that the male partner prefers cohabitation and female partner prefers marriage.

$$\begin{aligned} V_{Ht}^{M}(s_t|M) &\geqslant V_{Ht}^{C}(s_t|M) \\ V_{Wt}^{M}(s_t|M) &\geqslant V_{Wt}^{C}(s_t|M) \end{aligned} \tag{2.21}$$

$$\begin{aligned} V_{Ht}^{C}(s_t|M) &\geqslant V_{Ht}^{M}(s_t|M) \\ V_{Wt}^{C}(s_t|M) &\geqslant V_{Wt}^{M}(s_t|M) \end{aligned} \tag{2.22}$$

Unlike divorce, cohabitation is not a credible threat, which requires the cooperation of the other side. In such a case, I assume the female partner is the one who initiates the negotiation[19]. The female partner will increase the weight of the male partner in a cohabiting arrangement and induce him to participate in cohabitation. At the same time, the female partner is also better off cohabiting. In other words, the female partner tries to find a new weight to satisfy constraint 2.22. If such a new weight can be found, the household will choose to cohabitate with the weight changed to the new one[20]. Otherwise, the household will remain married.

Following the steps above, after a feasible and optimal living arrangement is chosen, $V_t(s_t|M)$, $V_{Ht}(s_t|M)$, and $V_{Wt}(s_t|M)$ can be determined. For a cohabiting household, the decision process is similar to the above. $V_t(s_t|C)$, $V_{Ht}(s_t|C)$, and $V_{Wt}(s_t|C)$ can be evaluated.

2.4.5.2 Divorced household

For the divorced household member, there is probability π that she gets married or cohabits again in each period, which is exogenous. With probability $1-\pi$, the household member remains divorced and maximizes individual utility. To make the model empirically trackable, I assume the household member expects to meet a potential partner the same as the previous partner in terms of observed characteristics. The unobserved matching qualities are drawn ran-

[19]In case of divorce, suppose husband wants to be married and wife wants to get divorced. Since divorce is a credible threat, husband will initiate the negotiation.

[20]To minimize the distortion, I assume household has the tendency to stay in the original living arrangement. Hence, at least one of the constrain in (21) has to be strictly satisfied,

domly. Let s_{it} be the relevant state variables for household member i at time t. If the divorced household member does not meet a potential partner in this period, she solves the problem below :

$$V_{it}^{D}(s_t|D) = max_{A_{it},cap_{it},imput_{it},h_{it}^{w},h_{it}^{hw}} u_{i,t}^{Divorce} + \alpha * E[V_{i,t+1}(s_{t+1}|s_t, D)] \quad (2.23)$$

subject to budget constraint for a divorced household member with no divorce cost involved, time constraint 2.11 and in house production constraint 2.7. If the divorced individual meets someone in the current period, she and the potential partner have the same status as cohabiting. They can then choose whether to cohabit, get married or not be together according to the decision process above. Thus, the expected utility[21] for household member i at time t is

$$V_{it}(s_t|D) = (1-\pi) * V_{it}^{D}(s_t|D) + \pi * V_{it}(s_t|C) \quad (2.24)$$

and the expected utility[22] at the end of period $t-1$ is

$$E(V_{i,t}(s_t|s_{t-1}, D)) = (1-\pi) * E(V_{it}^{D}(s_t|s_{t-1}, D)) + \pi * E(V_{it}(s_t|s_{t-1}, C)) \quad (2.25)$$

2.4.6 Estimation

I solve the model using backward induction. In this model, unions are considered formed when men are 25 years old. All individuals retire at age 65,

[21]The expectation is over whether the potential partner can be met

[22]The expectation is over state variables.

and die at age 75. In the last period, they spend all the assets with no bequest motive. To mitigate potential problems from this assumption, I only consider households with a male partner between 25 years old and 55 years old in 2002. There are six parameters to be estimated for each education type of households: mean and variance of unobserved matching qualities for both marriage and cohabitation, parameter for utility from leisure, and parameter for penalty if men do not work[23]. Remaining parameters are borrowed from literature or estimated using the reduced form analysis. Table E.4 shows lists of preset parameters. Wage process, decision on capital, re-marriage rate and re-cohabiting rate are estimated separately based on data. To estimate parameters for unobserved matching qualities, I consider these four moments: movement from marriage to cohabitation, movement from marriage to divorce, movement from cohabitation to marriage, and movement from cohabitation to separation. I consider female housework and male labor supply to estimate parameters for utility from leisure and penalty for men not working.

2.5 Results

Results for estimation are presented in Table E.6. Figure F.7 shows matching between Mexican data and simulated data of living arrangement movement for

[23]The mean and variance for unobserved matching qualities are education match specific. Parameter for utility from leisure, and parameter for disutility if men do not work are the same for two education match types.

low education households. Given that households are married in the current period, Figure F.7a shows the distribution of the living arrangement for these households in the next panel for low education households. The colored bar shows distribution in the data, and the empty bar shows distribution in the simulated data. Figure F.7b shows the results for cohabiting low education households. Figure F.8 shows matching of living arrangement movement for high education households. Table E.5 shows the corresponding percentage. Figure F.9 shows matching for male labor supply and female housework. For male labor supply, the interpretation for 8 to 10 hours is working 8 hours or more, but less than 10 hours. Similarly, for female housework, the interpretation for 0 to 4 hours is working 0 hours or more, but less than 4 hours. Remaining categories can be interpreted in similar ways.

Focusing on the living arrangement, the match of movement across living arrangements is better for married households than for cohabiting households. One discrepancy for cohabiting households is that there are more households getting separated rather than married in simulated data compared with real data. There are two potential explanations for such a mismatch. Firstly, the unobserved match qualities could be correlated within households. In each period within one household, the shocks of unobserved matching qualities for marriage and cohabitation could be correlated. For example, one explanation for a negative shock of cohabitation could be that a household wants to move to marriage, and is no longer satisfied with cohabitation. Hence, they draw negative shock for cohabitation and positive shock for marriage, and transit into marriage. However, as I model these two shocks independently, the model cannot capture such movement and will therefore predict more movement from cohabitation to separation. Secondly, I model that persistence for unobserved matching qualities only exists within the same living arrangement. In my current model, when a household moves from marriage to cohabitation, or from cohabitation to marriage, I assume the household loses all the previous accumulated unobserved matching qualities. This could be considered as a reasonable assumption for a household moving from marriage to cohabitation. For household moves from cohabitation to marriage, this assumption could be too strong. Most likely, a cohabiting household will also bring along part of the previous cumulated matching qualities when it moves to marriage. Without considering such a possibility, marriage will be less attractive for cohabiting households and more separation will be predicted based on the current model. The main concern for not modeling in these two ways is a lack of empirical moment for identification. For the first case, the correlation between two shocks needs to be estimated and for the second case, the proportion of unobserved matching qualities that can be transferred when households move between marriage and cohabitation needs to be estimated. Moreover, the true model could also be the combination of these two explanations.

In the sample, 90% of divorces and separations are triggered by men unilaterally, and the remaining 10% are triggered by both men and women mutually based on simulation. Individuals will trigger divorce and separation if the utility from the current union with highest power assigned is lower than the utility from living alone. To explain why men are more likely to trigger divorce, I examine the outside options for all the individuals in the sample as shown in Figure F.10. Figure F.10*a* shows female outside options, or the estimated value when households separate or get divorced[24]. Figure F.10*b* shows male outside options. To generate the graph, I calculate the value for each individual in case of divorce or separation from their current union. Then, I generate mean within each age interval for each household's type, based on current living arrangement and education match. On average, men have better outside options than do women, which provides one explanation for men being more likely to trigger divorce or separation. Individuals from high education households have better outside options compared with individuals from low education households. Married women have better outside options than do cohabiting women, in both high education and low education house-

[24]The results for 60 years old are noisy. The sample size is very small here and could be the explanation for noisy results.

holds. Cohabiting men have better outside options than do married men for low education households. For high education households, the comparison is less clear. However, if we average over all age groups, cohabiting men still have better outside options than married men. Two factors here can explain the differences of outside options between cohabiting individuals and married individuals: different demographic characteristics and laws related to asset division. In a later section, I will illustrate to what extent the difference in outside options can be accounted for by different asset division rules. One important point here is that men have better outside options, especially cohabiting men. This is one of the channels to explain the differences between the stability of married households and cohabiting households.

Figure F.11 shows gains from forming unions for low education households and high education households respectively, based on the sample[25]. Considering low education households, each line represents the average gain for a gender and living arrangement group. The gain is defined as the difference between the value from forming a union and from being single, either getting divorced or separated. The figure shows a snapshot of the gain for the Mexican sample. In married couples, young women gain more from forming a union than do

[25]The gain is at each specific age, not discounted to one age. As we are considering limited commitment model, households will consider the gain during each period. Hence, any policy should consider the gain for household at each period, not a discounted gain to one specific period, for example the first period.

young men. Such a difference disappears for older couples. Women gain more from forming a union at an earlier stage of life and men gain less. Over time, the gain from forming a union converges for women and men. One explanation for this convergence of gain from marriage is because of convergence of outside options. At an earlier stage of life, men have higher outside options compared with women. Over time, the outside options of men and women converge. A similar trend of convergence of gain between men and women from forming a union can also be observed in cohabiting households. One implication here is that over time, men become more attached to the union. As limited commitment model emphasizes on participation and incentive compatibility for each period, policy intervention may have different impacts on households along age dimensions; for example, when a policy is implemented to improve female welfare within households. This policy could be more effective in older households. For younger households, such a policy may violate the participation constraint for men and simply trigger renegotiation of bargaining power within households, or make forming a union less attractive.

Another observation based on Figure F.11 is that at earlier stage of life, lifetime gain from cohabitation and marriage are very similar for men and women respectively. Over time, the gain from these two types of union diverges, with the gain from marriage becoming greater with age. Hence, younger individuals choosing between marriage and cohabitation may get similar benefits from these two living arrangements. However, over time, due to different matching qualities and their accumulation, the gain from marriage becomes higher. Similarly, due to the limited commitment for each period, marriage becomes more stable over time compared with cohabitation. The aforementioned results hold true for both high education and low education households. This provides another explanation for the difference between the stability of married and cohabiting households.

Since the above results are from snapshots of the sample, one may interpret the previous results are due to cohort differences. The demographics and the sample size could also be different for different age groups and household types. To make interpretation of results easier, I consider a representative median household[26] from the low education group and one from the high education group. Figure F.13*a* shows gains from forming a union for these two households if they are married at age 25, while Figure F.13*b* shows gains from forming a union if these two households are cohabiting at age 25. Firstly, if we compare two living arrangements for each representative household, we can still obtain similar conclusions as using a snapshot of sample, as shown in Figure F.12. Over time, the gain for men and women converges within each

[26] A median household is defined as having wage and asset at the median among all households. There are two children. Male and female each owns 50 percent of asset.

living arrangement, and the gain between marriage and cohabitation diverges. Secondly, referring to Figure F.13, comparing two households across education groups, and within each living arrangement, men from high education households gain more from forming a union than do men from low education households. The difference between gains is greater in marriage compared with cohabitation. For women, the gain from marriage is greater for women from high education households, except at age 30. The gain from cohabitation is similar for women from both types of households. Over time, married high education couples gain more from forming a union compared with married low education couples . However, such a pattern is not observed within cohabitation. One explanation could be that high education households invest more in marriage than do low education households. However, these two types of households invest similarly when cohabiting, which could be due to a lack of intertemporal commitment.

To sum up, the above results suggest that men have better outside options compared with women. Men are more likely to trigger divorce or separation for both types of unions. Married men have lower outside options than do cohabiting men. Since the law for asset division in divorce favors women, this provides one explanation for marriage being more stable than cohabitation. Over time, individuals gain more from marriage than from cohabitation. As participation constraints need to be satisfied for every period, married couples are more likely to satisfy constraints. Both differences in matching qualities and demographic characteristics contribute to such divergences. Next, I consider three counterfactuals to examine to what extent unobserved matching qualities, asset division law, and difference in demographic characteristic differences explain the differences between the divorce rate for married households and the separation rate for cohabiting households.

In the first counterfactual, I change matching qualities for cohabiting households in the same way as married households. In the second counterfactual, I assume cohabiting households follow the same asset division rule as married households when they separate, in that female gets 70% of assets. In the third counterfactual, I combine both changes above. I calculate how the simulated separation rate for cohabiting households changes based on the data.

If cohabiting households have the same matching quality process as married households, the simulated separation rate decreases from 13.19% to 5.33 %, 59.6% decrease, for low education households. For high education households, the simulated separation rate decreases from 12.24% to 3.06%, 75.0% decrease. For both low education and high education households, changing unobserved matching qualities decreases the separation rate for cohabiting households by more than 50%.

If cohabiting households follow the same asset division rule as married households when a union dissolves, which means the cohabiting female also gets 70% of assets, cohabiting females are expected to gain and cohabiting males are expected to lose. Simulated data suggests that the average difference between the cohabiting female's outside options and the married female's outside options is reduced by 79.2% for low education households, and 49.1% for high education households. Previous analysis suggests that cohabiting women have lower outside options than do married women. Married women's outside options stay the same. Such a change is driven by improvement of cohabiting women's outside options. The number above also shows the extent of differences in outside options that can be explained by the law of asset division between cohabiting women and married women. One important rationale for such a policy is to improve women's utility upon separation from cohabiting unions. The average difference between cohabiting men's outside options and married men's outside options shrinks by 73.4% for low education households, and by 96.6% percent for high education households. On average, cohabiting men have better outside options than married men. Similar to married women, married men's outside options stay the same in this counterfactual. Cohabiting men's outside options decrease. This is a potential channel to explain the decrease in separation rate for cohabiting households through such regulation.

Since men are more likely to trigger separation, by decreasing their outside options, separating from their partners is less attractive. The simulated results suggest, with the same asset division rule imposed on cohabiting households, for low education households, the separation rate decreases from 13.19% to 11.56%, a 12.4% decrease. For high education match households, separation rate decreases from 12.24% to 11.73%, a 4.2% decrease.

In the third counterfactual, the separation rate for low education households decreases from 13.19% to 4.59%, a 65.2% decrease. For high education households, the separation rate decreases from 12.24% to 3.06%, a 75.9% decrease. Comparing cohabiting households with married households in the current counterfactual, the only differences now are the observed characteristics. These explain the remaining differences between the separation rate of cohabiting households and the divorce rate of married households in the sample.

Hence, for low education households, unobserved matching qualities explain 72.64% of the difference between the separation rate of cohabiting households and divorce rate of married households. Divorce law explains 15.06%. Jointly, these two factors explain 79.48% of differences. For high education households, unobserved matching qualities explain 93.96% of the difference between the separation rate of cohabiting households and divorce rate of married households. Divorce law explains 5.22%. Jointly, these two factors explain 93.96%.

The unobserved matching qualities explain the majority of the difference between the separation rate of cohabiting households and divorce rate of married households. The effect is much higher for high education households. As mentioned previously, both self-selection and different in-house investment contribute to differences in matching qualities. From the point of view of different in-house investment, this result could support literature suggesting the increasing return from in house investment along income and education dimension. Marriage can facilitate in-house investment. Such investment is more productive in high education and income households. For example, return from investment in children is much higher for high education and income households. Such a phenomenon may also exist along other dimensions of investment. Many tangible in house investments may show discrete natural, and low income households may not be able to purchase these huge assets. High education individuals may manage in-house bonding activities much better than do low education individuals. Potentially, high income and education households could gain more from forming unions. Hence, unobserved matching qualities may play more of a role in differentiating cohabiting households and married households for high education households.

The law of asset division explains around 15.06% and 5.22% of the differences between the separation rate of cohabiting households and divorce rate of married households for low education and high education households, respectively. This shows a short-run effect of such a regulation, as all the other factors stay the same. However, in the long run, the effect could be much larger depending to what extent the difference in matching qualities can be explained by different in-house investment and self-selection, respectively. If self-selection explains most part of the differences in matching qualities, such a regulation on cohabiting households will have little impact on improving matching qualities for existing couples. Otherwise, with additional regulation for cohabiting households, cohabiting couples may invest more in family bonding and public goods. The unobserved matching qualities for cohabiting households may converge to married households. As unobserved matching qualities have a much larger effect on decreasing the separation rate, the aim of such policies should be to improve matching qualities of couples, rather than to decrease outside options of individuals.

For future couples, the consequence of such a regulation would be more complicated. One important issue to consider is how such a policy affects the selection process into different unions. One concern would be that couples that can cohabit with no regulation may choose not to form a union in light of the additional regulation. Especially for men, forming a union may not be attractive compared with living alone. Women may get hurt in such a case as they gain more from forming the union. Without modeling the union formation, it is unclear how such a regulation will affect future generations.

2.6 Conclusion

In this paper, I examined factors affecting households' stability by adopting a dynamic household collective model with limited commitment. I considered three factors: unobserved matching quality, asset division which affects outside options, and different demographic characteristics. Specifically, I compared married and cohabiting households and examined to what extent the differences in stability between these two types of living arrangement can be explained by each factor. As suggested by literature, there is an education gradient for living arrangement preferences. I separated households into two groups based on education level.

Unobserved matching qualities explain most differences between the stability of married households and cohabiting households, and explain more for high education households than low education households. Differences between unobserved matching qualities could be due to the selection process at the initial stage when the union is formed. It could also be the results of different investments within households after a union is formed. Current work cannot separate effects from these two aspects. This could be studied in the future. For example, a framework with an initial matching process could be incorporated. From the literature, it was also found that high education and income households get a much higher return from their in-house investments, for example on children. This could be one explanation for the fact that unobserved matching qualities matter more for high education households. Due to the stronger intertemporal commitment generated within marriage, married households invest more within their households than do corresponding cohabiting households. Return from such investments could be higher for high education households compared with low education households. Hence, such unobserved matching qualities would matter more for high education households. On the other hand, it is also possible that high education individuals spend more time searching for suitable partners, which leads to better initial matching qualities. As mentioned before, studies considering the initial matching process should be performed to confirm this possibility.

Asset division explains 5% to 15% of differences in stability between married households and cohabiting households depending on education level, with a smaller effect on high education households. For married households, divorce law usually favors women by allocating more assets to them. By adopting a similar law for cohabiting households in the counterfactual, on the one hand, it can increase women's welfare upon separation from their current partners. One the other hand, it can decrease men's outside options, or utility upon separation. Both are confirmed by the simulated results. It is the second channel that can potentially decrease the separation rate for cohabiting households. In general, men tend to have better outside options, and are more likely to trigger separation. By decreasing men's outside options, they are less likely to trigger separation. From the perspective of couple as a whole, there is a possibility that dissolution of the current union is ideal. If matching qualities for the couple are low, it is not obvious whether it is ideal to force them to stay together by imposing additional laws. However, one important issue not calculated in this work is the welfare of children. Presumably, upon separation of a couple, children will suffer. And children from a cohabiting union are expected to suffer more. Further work should also be done to consider welfare changes for children.

Results of the counterfactual of asset division for cohabiting households could be interpreted as a short run effect. In the long run, the effect could be larger as it may change investment behavior within households for the cohabiting couple. Moreover, it can also alter the initial matching qualities for future couples. As argued in this paper, polices should be designed to help improve matching qualities of couples, rather than decrease individuals' outside options, if policy makers aim to increase stability of unions.

With cohabitation becoming more popular, works examining welfare implications of such a trend are relevant. This work examines this issue from the perspective of household stability. Future work from other perspectives, such as focusing on welfare distribution within a cohabiting union, children's welfare, and transition between cohabitation and marriage, are also important. Mexico, Latin American countries, and some north European countries are unique in the sense that cohabitation has a long history. These countries could serve as future reference for countries in which cohabitation is not yet common. This paper only considers asset division. Other regulations should also be examined to have better picture of social welfare implications; for example, alimony transfer.

Chapter 3

Ethanol-Blended Gasoline Policy and Ozone Pollution in Sao Paulo

3.1 Introduction

In recent decades, effective emissions control policies, particularly of volatile organic compounds (VOC), may have shifted the ozone production regime in urban atmospheres in the United States and elsewhere, from the "VOC-saturated" to the "VOC-limited" regime. Under such atmospheric conditions, further reductions in the volume and reactivity of VOC emissions can reduce the formation of ozone, a secondary pollutant that at the ground level harms human health, whereas reductions in nitrogen oxides (NOx) emissions can actually cause ozone concentrations to rise (Jacob 1999). Conversely, under VOC-saturated ("VOC-insensitive" or "NOx-limited") conditions, ozone production is insensitive to variation in VOC emissions and increasing in NOx emissions.

Knowledge of the atmospheric ozone production regime, and how it may be changing over time, is therefore critical to designing successful ozone control strategies. Such knowledge is not acquired without cost, however, as evidenced by a large atmospheric modeling literature that studies the chemistry of atmospheres over specific metropolitan areas or regions, and how its composition evolves over the annual, weekly and daily cycles. Models build on our current understanding of chemical and physical reactions of species in the atmosphere, and are based on often incomplete inventories of anthropogenic and biogenic emissions, by chemical species over time and in space. These computationally intensive differential-equation models are then used to predict how specific emissions control policies, for example, changes in the mix or composition of fuels utilized in the urban transport fleet, impact ambient pollutant concentrations, such as ozone (O_3) and particulate matter (PM). Madronich (2014) states: "uncertainties in numerical models often preclude firm conclusions regarding the dominant control over ozone chemistry in a city" (p.397). The difficulty in diagnosing the state of the local atmosphere is also suggested by the fact that, whereas decades of pollution control have led to the abatement of urban pollutants such as carbon monoxide (CO), NOx and PM, the record on ozone abatement is more mixed.

A case in point is the metropolis of Sao Paulo, in southeastern Brazil. Home to 20 million people, the Sao Paulo metropolitan area combines 40 municipalities and accounts for one-fifth of the country's GDP. Over the 1980s and 1990s, the metropolis experienced a steady process of deindustrialization, as factories moved inland or out of state. The economy today is dominated by services, and power generation is mostly hydroelectric. With a mild climate that requires minimal winter heating, the predominant source of anthropogenic emissions of ozone precursors—NOx, VOC and CO—is road transport. The circulating fleet is comprised of 6 million passenger cars, burning a mix of gasoline and ethanol that has varied over time, 1 million motorcycles predominantly powered by gasoline, and 120,000 diesel buses and trucks. Both gasoline fuel—in fact a blend containing a 20 to 25% volumetric component of pure ethanol, thus referred to as E20/E25—and ethanol fuel—in the form of E100—are universally available at the pump, one nozzle typically alongside the other. Among regulated air pollutants, ozone often exceeds the standard (CETESB 2013) and is the one species that did not exhibit a downward trend in the period 2000 to 2013 (Perez-Martinez et al. 2015).[1]

[1] Sao Paulo's maximum daily 8-hour ozone standard is 140 μg/m^3, or about 71 parts per billion (ppb). Sao Paulo State Decree 59113, issued April 23, 2013, signals a long-run reduction to 100 μg/m^3 (51 ppb), citing current scientific knowledge of the health damage of ozone. For comparison, the new (2015) US NAAQS for ozone is 70 ppb. The US EPA summarizes trends for the US as follows: "Nationally, average ozone levels declined in the 1980s, leveled off in the 1990s, and showed a notable decline after 2002" (epa.gov/airtrends/ozone.html, accessed on May 10, 2016).

An atmospheric modeling study, calibrated to Sao Paulo using the available emissions data, predicted that if all passenger cars were to run on ethanol (E100) rather than a combination of the ethanol and blended gasoline (E20/E25) fuels that heterogeneous consumers purchase at retail, then ozone levels would fall by up to 55% (Martins and Andrade 2008a). However, a recent empirical (econometric) study found the opposite effect (Salvo and Geiger 2014). The study exploited drivers of "flexible fuel" vehicles (FFVs) substituting between blended gasoline and ethanol as relative prices fluctuated at the pump, and concluded that substitution into ethanol caused ozone levels to *rise*. This single test suggested that modern-day Sao Paulo's atmosphere is VOC-limited, by which higher ozone levels associated with an increased ethanol fraction in the fuel mix was due either to (1) higher reactivity of VOCs from ethanol relative to gasoline use, whether via combustion or evaporation processes, and/or (2) lower NOx emissions from ethanol combustion relative to gasoline.[2] Salvo and Geiger's empirical finding was consistent with a controversial atmospheric modeling study that predicted that ozone levels would rise in Los Angeles, and elsewhere in the US, were passenger cars to adopt E85 ethanol over E10 gasoline (respectively, ethanol and gasoline fuels that are retailed in the US)

[2]Section 2 cites vehicle emissions standards legislation that allows carmakers to exclude (i.e., simply ignore) the mass of unburned ethanol, itself a reactive VOC, that is emitted from the tailpipe.

(Jacobson 2007, Ginnebaugh et al. 2010).

This paper, based on four separate quasi-experiments, provides an important second empirical test of the combined hypotheses that the Sao Paulo megacity's atmosphere is VOC-limited and that higher fractions of ethanol at the expense of gasoline in the light-vehicle fuel mix cause ozone levels to rise. Rather than take as identifying variation FFV owners choosing either E20/E25 gasoline or E100 ethanol at the pump as relative prices vary, the present study exploits discontinuities in the composition of gasoline fuel, between E20 and E25, mandated by the central government. We adopt a regression discontinuity design (RDD), with time as the forcing variable, and separately examine four discontinuities that occurred in recent years: out of E25 and into E20 in February 2010, back to E25 in April 2010, again into E20 in October 2011, and back to E25 in May 2013. On each occasion, the ethanol discontinuity in the volume of gasoline fuel sold at retail, of 5 percentage points, was not large. This design feature, coupled with the natural variability in ozone levels, suggests that the power of the test might not be high. On the other hand, the limited shift in fuel composition that the policy mandated—i.e., 5 percent of gasoline fuel—is an advantage of the design, since fuel changes were *not* salient to gasoline consumers, some of whom might otherwise have responded along the extensive margin of fuel choice (substitution with E100) or the intensive margin of driving (usage), potentially confounding our inference.[3] Moreover, it is fortunate for the purpose of our empirical test that, at the time of each of the four discontinuities, E20/E25 gasoline users accounted for a majority share, in terms of distance traveled, of light vehicles including motorcycles.[4]

Our results weigh in favor of the hypothesis that Sao Paulo's current atmosphere is VOC-limited and that raising the ethanol fraction in the fuel mix causes ozone levels to rise. Our preferred point estimates, based on local linear regression, suggest that ozone levels rose by 7-9% when E25 substituted for E20 (with standard errors of 2 to 3 log points), and these estimates are consistent with what Salvo and Geiger found using identifying variation of different design. Thus, our empirical finding joins the single empirical study to date in countering modeling work that suggested that ethanol use would abate ozone levels in Sao Paulo. As we argue below, to the extent that relative atmospheric concentrations of pollutants in Sao Paulo, interacted with its meteorology, reasonably resemble conditions found in some urban areas and seasons in the US today, our results suggest that ozone levels could rise as the ethanol fraction in gasoline grows beyond 10%, e.g., from E10 to E20, to meet statutory Renewable Fuel Standard (RFS) requirements (36 billion gallons of

[3]We also verify that prices per liter at the pump were not discontinuous at the four cutoffs.

[4]A parallel can be made with Auffhammer and Kellogg (2011)'s study, also based on RDD, of the effect on ozone pollution of policies shifting US gasoline content; fortunately, they do find that their tests have power.

biofuel by 2022, up fourfold since 2008).[5]

In contrast to ozone, we fail to detect changes in PM2.5 concentrations as the composition of gasoline shifted between E20 and E25. To the best of our knowledge, ours is the first empirical test of the relationship between gasoline-ethanol use and ambient PM2.5 levels (particulate matter with aerodynamic diameter up to 2.5 μm). Our findings for ozone and PM2.5 guard against statements, including those by industry-funded groups and subsidy-seeking politicians, that advocate ethanol on air quality or public health grounds.[6]

The paper makes three main contributions. First, it provides significant empirical evidence on the causal relationship between ethanol use and ground-level ozone in Sao Paulo (population 20 million). It detects a positive ethanol-ozone response by employing credible and different quasi-experimental variation—regulatory discontinuities in gasoline content—compared to Salvo and Geiger (2014), where consumers' response to prices also had to be estimated. Our result, as does Salvo and Geiger's, counters predictions for Sao Paulo based

[5]Arguing against the 10% "blend wall" and "arcane barriers" in the form of gasoline vapor pressure regulations, the US ethanol trade association claims: "research...demonstrates that gasoline blends containing 20-40% ethanol can deliver the octane needed to maximize efficiency in advanced internal combustion engines" (p.8, RFA 2016).

[6]For example, in a radio interview in 2010, a former Secretary of the Environment claimed that "last time the government mandated a reduction in the ethanol fraction in gasoline from E25 to E22, the environmental authority's air quality monitors, which are scattered throughout the city, immediately indicated a deterioration in air quality... and that was only with a 3 percentage point reduction" (CBN Noticias 2010). In the US, the American *Lung* Association of the Upper Midwest asserts that "E85 has been recognized as a Clean Air Choice" (cleanairchoice.org/fuels/e85.cfm, accessed on May 10, 2016).

on atmospheric modeling of a hypothetical shift in the fuel mix.[7]

Second, our paper enables subsequent research on the health effects of ozone. Deschenes et al. (2012) describe the "contentious current academic and policy debates about ambient ozone pollution" (p.3), in particular, surrounding the US EPA's recent tightening of ambient ozone standards.[8] By using a different design to validate a recently established empirical result, we aim to make shifts in the fuel mix a credible instrumental variable for ozone variation in the analysis of health outcomes, including hospitalizations and mortality. Research into how ozone in Sao Paulo affects public health lies beyond the scope of this paper and is being pursued separately. By helping to inform on ozone's health damage, our result has policy relevance beyond regions that share Sao Paulo's atmospheric conditions. Third, by targeting an economic policy audience and bringing added evidence to bear on what ozone chemistry theory indicates is an empirical question, the paper highlights a result that is relevant beyond Sao Paulo, such as US regions that share features of Sao Paulo's atmosphere and where policymakers are considering regulation to raise the ethanol content in gasoline. We further discuss ozone produc-

[7]Madronich (2014) argues that "a purely empirical approach (can) circumvent the problems associated with atmospheric chemistry modeling," adding that "empirical analysis...(is) needed to evaluate the reliability of atmospheric chemistry models designed to simulate the effects of the transportation sector on air quality" (pp.395-7).

[8]Deschenes et al. (2012), citing ten references in the academic and policy literatures, add: "These ozone standards are so contentious partly because there is substantial uncertainty about how ozone affects health" (p.3).

tion regimes in Section 2 and cite literature arguing that, following decades of VOC emissions control programs (US EPA 2004, Table 1), VOC-limited conditions similar to Sao Paulo's might be increasingly relevant in the US, including locations where ozone concentrations are not lower on weekends relative to weekdays (Stedman 2004, Stephens et al. 2008, Fujita et al. 2016). Chicago provides one such example where (1) recent research argues that "O_3 production in Chicago became more sensitive to VOCs starting in 2008/2009 and may have switched from being NOx-limited to VOC-limited" (Jing et al. 2014), and (2) higher ethanol penetration in the fuel mix in the form of an "E15 gas station ordinance" is being debated (e.g., Chicago Tribune 2015).

The balance of the paper is as follows. Section 2 lays out the hypothesis and discusses its potential relevance to the US. Section 3 presents key features of the policy setting and the data. Sections 4 to 5 describe the design we adopt, our specification, and our results. Section 6 concludes.

3.2 Hypothesis and increased relevance to the United States

To develop the hypothesis that increases in the ethanol fraction of blended gasoline raise ozone pollution in urban centers such as Sao Paulo, we (1) discuss the relevant ozone chemistry, including its relevance to US cities. We then briefly (2) report on tailpipe emissions (and associated) studies of gasoline and ethanol combustion, focusing on ozone precursors, and (3) describe the variation used in the single empirical test of a positive ethanol-ozone response to date.

3.2.1 Cross-disciplinary theory

Jacob (1999, Ch.12) provides a textbook treatment of photochemical ozone formation in the troposphere, whereby sunlight triggers reactions involving the interaction of nitrogen oxides (NOx = NO + NO_2) and VOCs (reactive hydrocarbons). These ozone precursors are emitted from local and regional anthropogenic and biogenic sources, such as vehicles, industry, crops and trees, and their relative abundance in the local (i.e., unmixed) atmosphere determines the ozone production regime. Two regimes, with ozone production characterized by a very different sensitivity to precursors, are described most simply by way of comparative statics; in particular, see Jacob (1999) Fig. 12-4 showing ozone concentration isoquants as a function of VOC and NOx emissions.

In a NOx-limited regime, say a rural area downwind of a city or suburb, ozone concentrations are increasing in NOx emissions and are insensitive to VOC emissions. In contrast, in a VOC-limited (VOC-sensitive) regime such as a large urban center, ozone concentrations are increasing in VOC emissions yet are decreasing in NOx emissions. Thus, formulating a strategy to abate ozone pollution requires knowledge of the regime. For example, in a VOC-limited environment, NOx control that is not accompanied by VOC control may allow ozone levels to rise. Alluding to the imprecision of emissions inventories that feed into computer simulations of atmospheric science, Jacob (1999) writes that "(t)he early models were in error in part because they underestimated emissions of hydrocarbons from automobiles, and in part because they did not account for natural emission of biogenic hydrocarbons from trees and crops" (p.238).[9]

Stedman (2004) provides a readable explanation (including to non-chemists) of time-dependent ozone production processes. Triggered by sunlight, reaction rates are increased at high temperature, e.g., in the afternoon hours, outside the colder months of the year. Stedman divides the physical time path of ozone—driven by wind and suppressed or produced by local precursors—into four distinct segments. In an "upwind segment" A, ozone typically flows in from sources that are upwind to an urban area, and even descends from the stratosphere. In a subsequent (time and space) "titration segment" B, having reached the urban area characterized by high economic activity, ozone blowing in from upwind is depleted by nitric oxide (NO) emissions, such as from road vehicles: "The more NO is emitted, the longer the ozone concentration level is

[9]For more on ozone chemistry, including the use of (estimated or assumed) emissions inventories as an input to atmospheric modeling studies, see Seinfeld and Pandis (1998) and Finlayson-Pitts and Pitts (2000).

suppressed. We expect to see ozone levels increasing sooner in time and closer to the urban area during weekends than during weekdays, because NO emissions are lower on weekends" (p.65).[10] Lower NOx emissions accompanied by higher ozone concentrations on weekends relative to weekdays, particularly on sunny summer afternoons, are observed in Sao Paulo. Stephens et al. (2008) references "the many locations throughout the world (where) (o)bservations of this effect have been made" (p.5318).[11] Of relevance to gasoline versus ethanol combustion, not only greater NO emissions but also lower VOC emissions increase the length (duration) of the ozone suppression segment.

Subsequent to the ozone-inhibiting titration segment B, say in the urban area in the early afternoon following high NO emissions during the morning commute, complex photochemistry in a "VOC-limited segment" C leads to the buildup of ozone. The higher the mass concentration and reactivity of VOCs, the faster ozone builds up. Thus, abating VOC emissions—limiting the ethanol fraction in gasoline being a possibility—has the twin benefit of lengthening the ozone suppression segment B and of reducing the rate of ozone buildup in the

[10]Under titration, NO reacts with O_3 (ozone) to produce NO_2 and oxygen. Thus, NO depletes ozone.

[11]For Sao Paulo, Salvo and Geiger (2014, Tables SV and SVI) report afternoon mean ambient ozone levels of 80 $\mu g/m^3$ on Sundays and public holidays compared to 68 $\mu g/m^3$ on non-holiday weekdays—divide by 1.97 to convert to ppb (at US STP). Stephens et al.'s (2008) list includes US cities/areas such as New York, Baltimore-Washington, Northern, Central and Southern California, Atlanta, Chicago, and Philadelphia. Fujita et al. (2016) predicts daily maximum 8-hour ozone concentrations in VOC-limited Los Angeles to rise 18% under 2030 baseline emissions—NOx and VOC emissions down 61% and 32%, respectively—relative to 2008, a situation referred to as "NOx disbenefit."

subsequent segment C.

Finally, Stedman (2004)'s account of ozone dynamics includes a fourth "NOx-limited segment" D that is only reached if VOC levels are sufficiently abundant relative to the amount of NOx input. As a consequence of VOC/hydrocarbon (including carbon monoxide, CO) emissions control in recent decades, Stedman explains that segments B (higher NO inhibiting ozone) and C (lower VOCs reducing photochemical ozone formation) are increasingly relevant across US metropolitan areas: "As hydrocarbon emissions have gone down nationwide in the United States, the rate of ozone buildup has decreased nationwide. As a result, even places like Los Angeles rarely, if ever, reach segment (D), the NOx-limited regime... The hydrocarbon-limited regime is lasting longer in time and larger in space as hydrocarbon emissions are reduced..." (p.66).

Critically, in an increasingly prevalent VOC-limited United States, increases to the ethanol content in US retailed gasoline that were to increase VOC or lower NO emissions might lead to rising ozone levels: "The important understanding of recent trends is that hydrocarbon concentrations and reactivity have gone down so much that even Riverside (well downwind of Los Angeles) hardly ever reaches this NOx-limited ozone segment before the sun sets... The continued trend to lower mobile source VOC emissions means that there are fewer places that benefit from NOx reduction. Most people live where ozone levels *increase* if NOx is reduced" (Stedman 2004, p.66).[12]

In sum, all other things being equal, fuel/engine combinations that raise VOC emissions and/or reduce NO emissions in a large urban area may increase the levels of ozone to which a population is exposed. Higher VOC and/or lower NO emissions are the likely mechanisms for ethanol consumption in urban Sao Paulo causing ozone levels to rise, as Salvo and Geiger (2014) found in a single study to date. In the setting they study, over one million bi-fuel vehicle owners shifted between gasoline and ethanol combustion as relative fuel prices varied markedly over time—the consumer price of ethanol (made from sugarcane) moving in tandem with the world price of sugar, and the price of gasoline fixed at a roughly constant level by the central government.

3.2.2 Tailpipe emissions, smog chambers, and atmospheric modeling

Unfortunately, one does not observe emissions from a large, representative circulating vehicle fleet—not only exhaust but also evaporative emissions—and operated under real-world conditions. These conditions include vehicle age,

[12]Among criteria air pollutants in the US, ozone has arguably been the trickiest to abate over the past decades. Stedman (2004) writes: "The observation of lower CO and hydrocarbon—ozone precursors—would lead all interested parties to expect continued reduction in ozone concentrations. Ozone concentrations however seem to be steady or even increasing in the last few years. This unexpected observation is partly explained by the discussion above which indicates that the improvement expected from the reduction in urban hydrocarbon and CO emissions has been offset by the concomitant reduction in NO emissions" (p.66). Also see Lin et al. (2000, 2001).

maintenance, engine setup including air-fuel ratio and temperature, driving behavior, powered by the fuels that are actually purchased by consumers, and so on. Attempting to fill this void, a large environmental engineering literature has developed, often reporting laboratory (chassis dynamometer) measurements of tailpipe emissions for a few vehicle/fuel combinations at a time. With regard to increasing the ethanol content in gasoline, or comparing ethanol- to gasoline-dominant blends, one reading of this emissions testing literature is that it is inconclusive, showing large variance for NO and the wide range of VOC species.[13]

Though NO emissions trends appear inconsistent, the significantly higher heat of combustion of gasoline compared to ethanol[14] suggests a potential for higher NO emissions from gasoline combustion. Salvo and Geiger (2014) discuss why one "may expect that more NOx will be produced during gasoline relative to ethanol combustion" (p.S11). For example, Hubbard et al. (2014) report tests for which "(e)missions of NOx decreased by approximately 50% as the ethanol fraction increased from E0 to E30-E40" but caution that because "(e)ngine calibration effects are manufacturer and model specific; emission

[13]See references in Salvo and Geiger (2014, Supplement Part B), and Wallington et al. (2016). Hochhauser and Schleyer (2014) write: "it is well known that the presence of ethanol in gasoline results in higher levels of evaporative emissions due to increased permeation of fuel through fuel system components such as elastomers and plastics" (p.3242).

[14]Hubbard et al. (2014) report 43 and 29 MJ/kg for gasoline- and ethanol-dominant E0 and E80, respectively.

trends... will *not* be the same for all FFVs" (p.861).

A more consistent pattern associates ethanol use with significant emissions of aldehydes, a class of VOCs with "strong potential for ozone formation" (p.512, Nogueira et al. 2014). Suarez-Bertoa et al. (2015) report an ozone-formation potential for E85 that is twice that of E5-E15. Even small changes in the gasoline ethanol content can have a large proportionate impact on aldehyde emissions. For example, Durbin et al. (2007) report that "(f)or acetaldehyde, a significant effect was found for ethanol, with an increase of 73% when ethanol increased from 0 to 10%" (p.4062). Consistent with ethanol's penetration in the light-vehicle fleet, studies of Sao Paulo's atmosphere report elevated concentrations of acetaldehyde. Studies include Martins and Andrade (2008a, 2008b), Martins et al. (2008), Orlando et al. (2010) and Nogueira et al. (2015). These studies also find high concentrations of atmospheric ethanol, i.e., emitted unburned out of the tailpipe or through evaporation along the ethanol supply chain. Tellingly, Brazilian Ministry of the Environment's Normative Instruction 54 of November 19, 2004 authorizes reported test results of vehicles powered by ethanol to exclude the unburned ethanol mass from reported VOC emissions.[15] In a recent field campaign in the US, de Gouw et al. (2012) find

[15]Specifically, Article 6 states: "On quantifying emissions of non-methane hydrocarbons from road vehicles powered by hydrated ethanol, the unburned ethanol component of emissions can be excluded." A personal conversation with an atmospheric scientist indicated that the neglected unburned ethanol mass and reactivity are likely significant.

that ethanol has become a ubiquitous compound in urban air, associating this development with the increased penetration of E10 gasoline blends. US fuel ethanol use as a proportion of gasoline use rose tenfold from 1% in 2000 to 10% by 2010.

Beyond tailpipe emissions tests, a few studies attempt to simulate the effect of gasoline and ethanol motor fuels on atmospheric chemical composition. In a study that was restricted to a smog chamber, Pereira et al. (2004) suggested that E100 ethanol use could raise ozone levels by 30% relative to E22-E24 gasoline. An atmospheric modeling study of the Los Angeles basin concluded that powering vehicles with E85 ethanol versus gasoline would raise ozone concentrations by a range of 7 to 39 ppb for the conditions studied, and also increase ambient levels of acetaldehyde, formaldehyde, and peroxyacetyl nitrate (Ginnebaugh et al. 2010).[16]

In contrast, a mathematical modeling study calibrated to Sao Paulo's atmospheric system predicted significantly *decreased* levels of ozone from adding ethanol to the transportation fuel mix (Martins and Andrade 2008a). A recent review article underscores the need for air quality studies that are based on observational (field) rather than simulated data (Anderson 2009).[17]

[16] In other modeling studies of the US, Cook et al. (2011) simulated that adoption of E85 and E10 would raise ozone levels, while Nopmongcol et al. (2011) predicted a negligible change from E85. Results are very sensitive to assumed tailpipe emissions, e.g., NOx, across fuels.

[17] Anderson (2009) writes "(f)or the most part, we ignore the hundreds of individual com-

3.2.3 Empirics on ethanol and ozone: A single test to date

E100 ethanol and E20/E25 gasoline are widely available to Sao Paulo's consumers, and relative prices for the two fuels have moved significantly—though smoothly—over time. Salvo and Geiger (2014) examined variation in the proportion of bi-fuel (flexible fuel) vehicle owners choosing ethanol over gasoline fuel at the pump, i.e., the "ethanol share in the bi-fuel fleet," as relative prices varied between late 2008 and mid 2011. Bi-fuel gasoline-ethanol engines accounted for about one-half of vehicle miles traveled (VMT) by passenger cars circulating in the Sao Paulo metropolis; dedicated (single-fuel) gasoline engines burning E20/E25 accounted for the other half of VMT.

The present study focuses on gasoline blend discontinuities—i.e., E20 versus E25—as a source of identifying variation, rather than consumers substituting between gasoline (E20/E25) and ethanol (E100). Before examining these blend discontinuities, it is useful to describe Salvo-Geiger's source of variation. Figure I.1(a) indicates how the price of one liter of ethanol evolved relative to the price of one liter of blended gasoline at the pump, which was quite stable; panel (b) plots the ethanol share in the bi-fuel fleet estimated from

pounds that are actually emitted from the vehicle in that broad range of VOC compounds. As fuel composition changes, it is necessary to look at the details of the VOCs and how they change with changing fuel composition. If we do not, there can be dramatic effects on air quality, as we have seen in Brazil" (p.1034).

these prices. Because Salvo-Geiger did not observe fuel shares, they imputed these from a first-step demand model estimated from observed fuel prices and a surveyed distribution of consumer and vehicle characteristics for Sao Paulo city (Salvo and Huse 2013). As we explain below, we flexibly use relative price data (a) to control for such variation. Controlling for the estimated ethanol share (b) would be almost equivalent, given the almost linear empirical relationship between shares and relative prices (Salvo and Huse 2013),[18] except that we would need to correct for sampling variation in generating (rather than observing) series (b).

In the Appendix, we use a longer sample than that used by Salvo-Geiger, 55 months from late 2008 to mid 2013, estimating a positive ethanol-ozone response in Sao Paulo. We also estimate two variants to their two-step empirical model of fuel shares and ambient ozone concentrations. Madronich (2014) discusses the possible chemical mechanism in support of the observational evidence on ozone, and cautions that "the observed reduction in ozone levels should not be taken as evidence that a switch from ethanol to gasoline would improve air quality overall... a switch from ethanol to gasoline probably stimulates the production of secondary organic aerosols" (p.397). Motivated by Madronich's perspective, we take advantage of the present study's regression

[18] To see this, note that in Figure I.1 series (b) is almost the mirror image of series (a).

discontinuity design and extended sample period to look for an effect of the ethanol fraction also on fine particles. To the best of our knowledge, this is the first observational study to examine the relationship between a change in the ethanol-gasoline fuel mix and ambient PM2.5 levels.[19]

3.3 Policy setting and data

We exploit discontinuities in the volumetric proportion of (anhydrous) ethanol that distributors are required to blend into gasoline to test whether ozone concentrations in Sao Paulo rise as the light-vehicle fuel mix shifts from gasoline to ethanol and, similarly, whether ozone falls when the mix shifts back to gasoline. Over the period between 2008 and 2014, the central government mandated four changes in ethanol-blended gasoline fuel—see Table H.1. During these years, unblended gasoline E0 was not available at retail, only in blended form, either E20 or E25, and only a single blend was dispensed at any given point in time. Blended gasoline was known to consumers simply as "gasoline." The mandates applied to all fuel sold by distributors to retailers across the country. Compliance was likely very high, in part because fuel dis-

[19]Madronich (2014) conjectures: "ozone is just one component of photochemical smog. Particulate matter, another key component, accounts for a large fraction of the health impacts. Urban particulate matter is largely composed of secondary organic aerosols, the product of photochemical reactions between volatile organic compounds, nitrogen oxides and ozone. Although yields of secondary organic aerosols rise with ozone concentrations, they also increase in the presence of heavier volatile organic compounds, such as those emitted by the combustion of gasoline" (p.397).

tribution was very concentrated among a few firms (even more so than fuel retail). These policy shifts and reversals in gasoline content were motivated by the administration's industrial policy, including bargaining with the sugar industry, and were not induced by air pollution in the Sao Paulo metropolis (Angelo 2012, Salvo and Huse 2011, 2013).

Labeled in chronological order, discontinuities # 1 and # 3, effective for distributor shipments beginning February 1, 2010 and October 1, 2011 respectively, each consisted of 5 percentage point (percent) decreases in the proportion of ethanol in blended gasoline, from E25 to E20, i.e., one-quarter to one-fifth by volume. Discontinuities # 2 and # 4, effective for distributor shipments beginning May 1, 2010 and May 1, 2013, each consisted of 5 percentage point increases in the ethanol fraction, from E20 back to E25.[20] For ease of exposition, we present all results as corresponding to an *increase* in the ethanol fraction. Thus, for discontinuity # 1, February 15 2010 would fall in the "5 percent less ethanol" period (E20), and January 15, 2010 would fall in the "5 percent more ethanol" period (E25).

[20]The shift from E25 to E20, effective from February 1 to April 30, 2010, was announced as temporary on January 11, 2010 (Ministry of Agriculture Ordinance 7). Discontinuity # 3, again from E25 to E20, was announced on August 31, 2011 (Ordinance 678-2011) as open-ended, as was discontinuity # 4 back to E25, announced on February 28, 2013 (Ordinance 105-2013). Announcing a change a month or two ahead of its implementation is intended to allow the supply chain to adjust. Thus, ethanol production, which is concentrated in northwestern Sao Paulo state at least 400 km from the metropolis, was unlikely to adjust at the date cutoffs. Moreover, in an email interview, a former head of the sugar industry trade association (UNICA) stated that distributors were unlikely to make the change ahead of the mandated deadline.

Over this period, there were no changes in the composition of (hydrated) ethanol sold to retailers, namely E100, known to consumers as "ethanol." The two retailed fuels, ethanol (E100) and blended gasoline (E20 or E25) were ubiquitously available for consumers to purchase at the pump. Weekly surveys of about 350 retail stations in Sao Paulo city indicate that distribution of both gasoline and ethanol remained essentially universal throughout our study period.[21]

As described in Section 2, about one-half of miles traveled by passenger cars in the Sao Paulo metropolis, with a fleet size of 6 million, were equipped with dedicated (single-fuel) gasoline engines. Bi-fuel vehicles accounted for the other half of VMT, and their drivers tended to substitute at the pump between gasoline (E20/E25) and ethanol (E100) as relative prices varied. Whereas an individual bi-fuel vehicle driver might substitute between gasoline and ethanol at a given relative price point, substitution in aggregate occurred smoothly over a wide range of price variation (Salvo and Huse 2013). This means that changes in fuel shares in the bi-fuel vehicle subpopulation can be captured by a flexible trend, as discussed in the next paragraph. Also circulating across the metropolis were, roughly, 1 million motorcycles. These motorcycles were predominantly equipped with single-fuel gasoline engines.[22]

[21] Pump-level prices and availability are from the National Agency for Oil, Biofuels and Natural Gas (ANP).

[22] About 120,000 diesel trucks and buses complete the road vehicle fleet (Salvo and Geiger

Whereas Salvo and Geiger (2014) examined smooth changes in the bi-fuel vehicle share across gasoline and ethanol, in this study we focus on the relatively narrow time window around each discrete, abrupt change in the gasoline blend, from E25 to E20 and back to E25 repeatedly, among consumers who used gasoline fuel. Gasoline users consisted of dedicated engines and a gradually varying share of bi-fuel engines, substituting gasoline for ethanol at the pump. We control for this gradual substitution between gasoline and ethanol among bi-fuel vehicles by employing a flexible polynomial as well as observed fuel prices. This is done separately by discontinuity, taking a window of plus and minus 90 days around each cutoff date, which we explain below. Figure I.2 shows that fuel prices at the pump (including diesel fuel used by heavy-duty vehicles) varied smoothly, if at all, in the neighborhood of each gasoline blend discontinuity. We also include other time-varying controls that move smoothly, such as traffic congestion to proxy for vehicle use.

As Table H.1 indicates based on available fuel quantity data, at the time of each discontinuity, blended gasoline fuel accounted for the majority share

2014). Diesel combustion is a major source of NOx and particles (He et al. 2016). These emissions contribute smoothly to "background" variation in pollution since the volume of diesel combustion did not jump at the blend discontinuities. Moreover, to the extent that road emissions in other cities (or highways) in the state affected, through atmospheric transport (Lin et al. 2000, Lin 2010), pollution in the metropolis, shifts in the fuel mix were similar. In particular, blending requirements applied throughout the country on the same dates, and gasoline and ethanol prices and usage were similar throughout the state (Figure I.1). Thus, changes in background pollutants are unlikely to confound our inference of the ethanol-ozone relationship. The largest city within 600 km of the metropolis has 3% of its population.

of light-vehicle fuel sold in the state of Sao Paulo. At the available level of aggregation, blended gasoline usage ranged from 53% of total VMT by the light-vehicle fleet around discontinuity # 2, to 76% of VMT around discontinuity # 4. From the fuel market regulator (ANP), we obtained monthly fuel shipments reported by distributors, separately for E20/E25 gasoline and for E100 ethanol, and converted millions of cubic meters into approximate aggregate VMT shares. While these reported quantities include supplies to the state's highway market, which can vary by season, the quantity data supports two points we made previously. First, that gasoline fuel dominated ethanol fuel at the pump on all four occasions, justifying why our research designs exploits 5-percent discontinuities in its composition. Second, the aggregate gasoline fuel share varies across discontinuities, e.g., higher for discontinuity # 4, lower for discontinuity # 2, again justifying why we choose to examine each discontinuity separately, despite the cost of statistical power.[23]

Though E20/E25 gasoline was the main source of combustion in the metropolis (CETESB 2013), one empirical challenge we face is that a 5 percentage point shift in the blend is not large, and we may lack statistical power. At the same time, a somewhat small shift in fuel composition, which was not salient to

[23]The blended gasoline share reported in Table H.1 is one minus the aggregate ethanol share shown in Figure I.1(c). Converting fuel quantities from m^3 to VMT requires assumptions on the fleet's fuel economy (see the notes to Table H.1), but using "barrels of oil equivalent" also reported by ANP yielded similar shares.

consumers,[24] is less likely to induce confounding changes in consumer behavior, whether along the extensive or intensive margins. An example of a confounding shift along the extensive margin would be bi-fuel vehicle drivers who might substitute between E100 ethanol and gasoline at the pump whenever the composition of gasoline shifts abruptly between E20 and E25. Confounding intensive margin changes would be driven by motorists who might adjust their vehicle usage as the composition of gasoline shifts. Consistent with Figure I.2, in the Appendix we use station-level data (and a regression discontinuity design) to show that the price per liter of gasoline did not significantly change (neither statistically nor economically) as its composition shifted by 5 percent. Thus, confounding shifts along the extensive and intensive margins of consumer choice are likely insignificant, or not a source of concern.[25] The change that matters at each discontinuity point is the regulator mandated shift in gasoline composition.

Another empirical challenge is that the shift in the gasoline blend might have materialized over a few days. Whereas the ordinances set a sharp dead-

[24]To emphasize, there were no changes in the labeling of gasoline at the pump as the blend shifted.

[25]The usable energy content in ethanol, a partially oxidized hydrocarbon, is about one-third less than that in E0 gasoline (already accounting for differences in combustion efficiency). Hence, fixing the price per liter, a gasoline blend change from E20 to E25 would in effect raise the price of the fuel in $ per km by about 1.7%. To the extent this effective price change is salient to some gasoline consumers and they respond by switching to a choice of ethanol at the pump, this would help us detect a change in ozone pollution from increased ethanol use.

line to distributors for changing the blend, and were announced at least several weeks ahead, to impact Sao Paulo's atmosphere each change had to work through downstream inventories, particularly in the tanks of cars and motorcycles circulating in the metropolis. Our preferred specification uses a sharp design, with the regression discontinuity falling three calendar days after the government's deadline to distributors, e.g., February 4 for the first discontinuity's shipment deadline of February 1, 2010 (Table H.1). This three-day lag is meant to account for downstream inventories of older fuel, in tanks at both retailers and consumer vehicles. Salvo and Huse (2013) found that the majority of consumers purchase less than half a vehicle's tank every time and thus stop to refuel every few days. This suggests that aggregate fuel inventories downstream of distributors should be low, providing some support for our sharp design with a three-day lag. Note that the blend change can additionally impact emissions via evaporation along the supply chain, prior to combustion at the end point of usage, and that there is no separate deadline for retailers.

It is also worth emphasizing that we choose to flexibly examine each of the four discontinuities separately. This comes at the potential cost of statistical power. The benefit is that this allows us to control for otherwise potentially important confounders of air pollution, including seasonality in meteorological conditions, and confounding variation in economic activity and anthropogenic emissions. This matters since the different discontinuities happened at different times of the year, such as the fall month of May versus summertime February, and in different years, e.g., 2013 versus 2010.

Data. We combine highly spatially and temporally resolved observations of pollutant concentrations and meteorological and road traffic conditions for the four 180-day windows, each with a blend discontinuity at the center. From the Sao Paulo State Environmental Protection Agency (CETESB), we obtained hourly mass concentrations, in $\mu g/m^3$, for O_3 and PM2.5 at all the EPA's air monitoring sites in the metropolitan area of Sao Paulo. Besides ozone, our main pollutant of interest, we consider fine particles as their concentration is now monitored (this was not the case in the sample period Salvo-Geiger considered) and the conjecture that ethanol use may impact PM2.5 levels differentially to gasoline (Madronich 2014). We also obtained concentrations for CO and NOx (and its separate components, NO and NO_2) in their measured units, parts per million (ppm) and parts per billion (ppb), respectively. Table H.2 reports summary statistics for the combined four 180-day samples.

Figure I.3 shows the location of the sites monitoring one or more pollutant.[26] During the sample period, there were many more O_3 monitoring sites

[26]Pictures of sites are available in the online appendix to Salvo and Geiger (2014), specifically, page A-S21 on.

than there were PM2.5 sites. Historically, ozone is the pollutant that has most often exceeded ambient air quality standards (CETESB 2011). Ozone exceedance episodes tend to occur outside of the winter months of June to September.[27] Fortunately, each of the four discontinuities we examine also happened outside of this brief winter period. Table H.2 illustrates ozone's reactivity: ambient O_3 concentrations in the afternoon hours are about triple those in the morning, when radiation and temperature are lower, i.e., 1-hour means (maximums) of 66 versus 23 (353 versus 179) μg/m^3, respectively. For perspective, Brazil's federal standards include a 1-hour mean ozone level of 160 μg/m^3, and the US EPA recently revised its 8-hour ozone standard to 70 ppb, or about 138 μg/m^3. Ozone levels also vary widely in space, with afternoon means ranging from 77-84 μg/m^3 for sites 31 and 5 to 55-56 μg/m^3 for sites 29 and 27 (not reported for brevity). Reflecting both the spatial variability and the public health risk, in 2012 the EPA increased the number of O_3 monitors in the metropolis by 40%.

Likely because the literature on health effects of fine-particle pollution is relatively recent (Dominici et al. 2014), PM2.5 monitoring began only in January 2011 and at a single site. PM2.5 monitoring had increased to only five

[27]Solar radiation and temperatures, and thus atmospheric ozone production, fall in winter relative to the remaining months of the year, but the winter is mild compared to winters in much of North America.

sites by 2013. In addition to O_3, as Figure I.3 indicates, there is widespread monitoring of CO and NOx. This is likely due to road transport being the main source of anthropogenic emissions in the Sao Paulo metropolitan area: emissions inventories published by the EPA consistently put: (i) passenger cars and motorcycles (powered by E20, E25 or E100) as accounting for the majority share of CO and VOC emissions across all sectors of economic activity, e.g., 91.1% and 70.8%, respectively, according to CETESB (2012); and (ii) heavy-duty vehicles (trucks and buses burning diesel) as accounting for the majority share of NOx emissions, i.e., 60.3% according to CETESB (2012). Beyond O_3, PM2.5, CO and NOx levels, VOCs are not automatically monitored by the EPA,[28] and SO_2 monitors have been gradually deactivated or moved inland as the Sao Paulo metropolis deindustrialized over the 1980s and 1990s (industrial emissions tend to be the main source of SO_2, CETESB 2012). Importantly, power generation in southeastern Brazil is mostly hydroelectric, and mild winters imply minimal residential heating.[29]

Several of the EPA's pollutant-monitoring sites double as weather stations measuring, also on an hourly basis, ground temperature, solar radiation, rela-

[28]Routine monitoring of VOCs is uncommon, presumably due to cost. For example, in a study of the Chicago area Jing et al. (2014) recommend "(i)ncreased attention should be paid to improving the quantification of VOC sources, enhancing the monitoring of reactive VOC concentrations" (p.630).

[29]For example, hydroelectric plants accounted for 14,226 MW out of a total 19,555 MW, or 73%, of the electricity generating capacity that was installed in the state of Sao Paulo in 2009 (Negri 2010).

tive humidity, and wind speed and direction. To these meteorological data we added hourly precipitation and daily hours of sunshine, recorded by the Institute for Meteorology (INMET) at a site in the metropolis. We also obtained readings, every 12 hours, of the presence and height of thermal inversions, as these interfere with the dispersion of pollutants (Arceo et al. 2016, He et al. 2016). Thermal inversions are monitored by the Brazilian Air Force (FAB) from one location in the city. Table H.2 shows the conditions we control for (for brevity, we omit evening meteorology).

To further control for variation in pollutant concentrations, we use data on road congestion from the city's traffic authority (Companhia de Engenharia de Trafego, CET). We observe, at 30-minute intervals, which parts of an extensive fixed grid, totaling 840 km of road corridors, were congested. Following He et al. (2016), we partition these roads into geocoded segments of average length 80 meters and observe, at every point in time, whether each segment was congested (i.e., experienced stop-and-go traffic) or not. We can then integrate over space to obtain a measure of road congestion that is local to each air monitor, e.g., within a 2 km radius, or aggregate across the 840-km grid for a measure of citywide road use.

3.4 Research design and specification

3.4.1 Regression Discontinuity

Regression discontinuity designs have recently gained popularity among researchers seeking to identify causal relationships. The method takes advantage of a discontinuity in treatment probability near a cutoff point in a running variable, to identify the treatment effect for a subgroup of the population. Potential confounding factors are assumed to change continuously around the cutoff point. An example of a potential confounder in our setting would be price-induced variation in gasoline versus ethanol choices by consumers driving bi-fuel vehicles. There are two classes of RDD: sharp and fuzzy. In sharp RDD, treatment is a deterministic function of the running variable, changing from a weight of 0 to 1 at the cutoff point. In fuzzy RDD, the probability of receiving the treatment need not change from 0 to 1 at the cutoff.

Hahn et al. (2001) develop formal identification assumptions for treatment effects in this framework. They propose local linear nonparametric regression techniques to avoid the poor boundary behavior of the kernel regression estimator. In general, the treatment effect, τ, is:

$$\tau = \frac{lim_{x \downarrow c} E[Y|X=x] - lim_{x \uparrow c} E[Y|X=x]}{lim_{x \downarrow c} E[T|X=x] - lim_{x \uparrow c} E[T|X=x]}$$

where Y is an outcome variable, X is a running variable that governs the treatment probability, and c is a cutoff point. $T_i(x)$ is the treatment status for individual i for x in some small neighborhood around c. $T_i(x)$ is equal to 1 if individual i receives treatment. In sharp RDD, the preceding expression is simply $\tau = lim_{x \downarrow c} E[Y|X = x] - lim_{x \uparrow c} E[Y|X = x]$.

In empirical implementations, researchers may choose to use data that, in terms of the running variable, are realized at some moderate distance from the cutoff. For example, sample size might be a concern. To reduce potential bias from including such observations, one may control for additional covariates. Imbens and Lemieux (2008) find that including covariates likely does not affect identification and may improve the precision of estimates.

Even in the presence of control variables, one might not feel confident that $lim_{x \downarrow c} E[Y|X = x]$ and $lim_{x \uparrow c} E[Y|X = x]$ can be well approximated by a global linear function. It is therefore common practice in RDD applications to specify a variety of high-order polynomial functions of the running variable with the hope of fitting the data on either sides of the cutoff point. The treatment effect is then obtained from the behavior of the polynomial tails near the cutoff point. Recently, Gelman and Imbens (2014) argued against such practice, since high-order polynomials can assign excessive weight to observations that are distant from the cutoff, conflicting with the idea of RDD. There are no clear guidelines for choosing the order of polynomial, and estimates can be sensitive to modeling choice. Gelman and Imbens recommend the use of estimators based on local linear or quadratic polynomials and other smooth functions.

Implementing local polynomial regression requires that the researcher specify the kernel function, polynomial order and the bandwidth. Standard kernels include the triangular, uniform and Epanechnikov kernels. In our work, we specify a triangular kernel function. Local polynomials are typically specified to be linear or quadratic. In our work, we adopt local linear regression. Perhaps the more controversial choice is specifying the bandwidth. The nonparametric literature suggests that estimation results can be sensitive to bandwidth selection (Calonico et al. 2014a). In an RDD framework, Ludwig and Miller (2007), DesJardins and McCall (2008), Imbens and Kalyanaraman (2012), and Calonico et al. (2014b) study bandwidth selection criteria, making different recommendations. In our work, we adopt bandwidth tests developed by Calonico et al. (2014b) and, alternatively, by Imbens and Kalyanaraman (2012), hereafter referred to as CCT and IK, respectively. These criteria are asymptotically optimal under square error loss.

RDD in the evaluation of environmental and energy policy. In the environmental economics literature, specifically with regard to the causal effect of policies on air pollution, four recent RDD applications are worth noting. Davis (2008) and Lin Lawell et al. (2016) examine the effect of policies restricting driving on urban air in Mexico City and Bogota. Auffhammer and Kellogg (2011) examine the effect of gasoline content regulation on air quality, ozone in particular, in the US. Chen et al., (2013) examines the effect of latitude-based heating subsidies on particle pollution and life expectancy in China. Our setup is similar to the first three studies in that time is the running variable, and their authors argue that the policies were implemented immediately and achieved near universal compliance. We argue that such assumptions similarly fit our setting. Similar to Auffhammer and Kellogg, our focus on ozone further makes RDD suitable, in that its chemical reactivity implies that ambient air concentrations change quickly, for example, over the diurnal cycle, or upon a step change before and after a policy comes into effect. Whereas in both Davis' and Auffhammer and Kellogg's settings regulation was introduced with an aim to curb external damage from air pollution, in our setting any effect on urban air from mandated changes to the gasoline blend was an unintended consequence. It is this unintended consequence of ethanol use that our study seeks to uncover. In Chen et al., the running variable is geographic location (north versus south of the Huai river), and a rise in particle pollution was the unintended consequence of an energy policy.

3.4.2 Specification

An observation in our study is an air monitoring site-hour-date triple. For example, to evaluate the effect on ozone from changing the blend mandate, a single observation would be the ozone monitor at Site 5 (Ibirapuera) at 3pm on February 15, 2010. Consider the regression model, to be implemented separately by discontinuity (for brevity we omit subscript d):

$$y_{iht} = \alpha_{i0} + \alpha * treat_t + f_{ih}(date_{iht} - date_c) + \beta_{ih} * W_{iht} + \epsilon_{iht} \tag{3.1}$$

The dependent variable, y_{iht}, is the natural logarithm of the measured pollutant mass concentration, in the recorded units of μg/m^3 or ppm, where i, h, and t index site (monitor), hour, and date, respectively. Binary variable $treat_t$ is equal to 0 if the gasoline blend combusted on date t is E20, and 1 if the gasoline blend is E25, indicating lower and higher ethanol fractions respectively. To repeat from Section 3, we normalize the reporting of results to represent the treatment effect of policy that increases the ethanol component, from E20 to E25.

The eighth-order polynomial in date, $f_{ih}(date_{iht} - date_c)$, is centered at the policy implementation date, $date_c$. The polynomial trend is site and hour specific, to flexibly capture seasonal and unobservable trends at the site-hour level over the half-year (180-day) sample. Similarly, the vector of controls, denoted W_{iht}, is allowed to affect or co-vary with pollutant concentrations differentially by site and hour pair, via the parameters β_{ih}.

Included in W_{iht} is a vector of day-of-week fixed effects, to account for (site-hour varying) weekly cycles, e.g., systematic differences between, say, Tuesday 3pm and Friday 3pm, in the volume of surrounding vehicle traffic impacting ozone levels at site 5. Day-of-week fixed effects also capture weekend and public holiday variation in pollutant concentrations.

Other determinants of pollutant concentrations included in W_{iht} are meteorological conditions, road congestion at one or more levels of proximity to the monitoring site, and the ethanol-to-gasoline price ratio at the pump, which smoothly drives the choice of ethanol over gasoline fuel in the subpopulation of bi-fuel vehicles. Specifically, we control for: a quadratic in log temperature (^{0}C, measured in the contemporaneous hour h); a quadratic in log radiation (W/m^2); a quadratic in log relative humidity (%), a quadratic in log wind speed (m/s); indicators for each of four wind direction quadrants (whenever wind blows in excess of 0.5 m/s); precipitation (mm, from h to $h-3$); and indicators for the base of an atmospheric thermal inversion layer recorded lying within 200m, or between 200 and 500m, from the ground. We control for the quantity of road users by taking, on top of site-hour specific day-of-week intercepts, the extension of traffic congestion across the city recorded from 7am to 11am on date t (interacted with an indicator for t being a non-public holiday weekday).[30] The ethanol-to-gasoline price ratio enters vector W_{iht} as a third-order polynomial (and we lag this by four days, as in Salvo and Geiger 2014). We selected these controls based on the sensitivity of pollutant concentrations to their variation, for example, ozone's sensitivity to radiation and temperature, or the sensitivity of pollutant concentrations to wind speed. These controls have added importance recalling that the effect on air of a 5-percent discontinuity in the gasoline blend may not be large and we may lack precision. To emphasize, controls are flexibly interacted with site by hour fixed effects.[31] Appendix Figures I.8 to I.11 summarize variation in meteorology and road congestion around each discontinuity.

The treatment effect is estimated in two steps. First, using ordinary least squares (OLS) we estimate a restricted version of equation (1), that excludes $treat_t$ (but includes the site-hour specific eighth-order polynomials in date),[32] generating fitted residuals of ln(pollutant concentration). Intuitively, this de-

[30] Given the opening on March 30, 2010 of the Greater Sao Paulo beltway, that removed trucks from specific inner city roads (He et al. 2016), we also add a dummy variable (interacted with site-hour) to indicate dates post inauguration; these may be relevant to the first two discontinuities, with cutoffs on February 4 and May 4, 2010.

[31] In a robustness test, we interact controls only with site, not site-hour, fixed effects (Appendix Table H.10). For comparison, using a sample over multiple years, Davis (2008) specifies day-of-week (weekend) by hour fixed effects, but he does not additionally interact these with monitoring sites in Mexico City.

[32] Again, these flexible polynomials account for unobserved smoothly varying trends over each 180-day sample. In robustness tests, we vary the order, e.g., seventh-order, as well as drop these trends from the specification.

means the data and partials out the effect of other factors. Second, we implement local linear regression using the fitted residuals to estimate the treatment effect.

Denote the residuals generated in the first step by $\tilde{y}$. Formally, the effect of the "ethanol policy," α, is estimated as:

$$(\hat{a}_{b+}, \hat{\gamma}_{b+}) =_{a,\gamma} {}_{i,t,h}\{(\tilde{y}_{iht} - a - \gamma(date_{iht} - date_c))^2 \tag{3.2}$$

$$* K_b(date_{iht} - date_c)\mathit{1}(date_{iht} \geq date_c)\}$$

$$(\hat{a}_{b-}, \hat{\gamma}_{b-}) =_{a,\gamma} {}_{i,t,h}\{(\tilde{y}_{iht} - a - \gamma(date_{iht} - date_c))^2 \tag{3.3}$$

$$* K_b(date_{iht} - date_c)\mathit{1}(date_{iht} < date_c)\}$$

$$\hat{\alpha} = \hat{a}_{b+} - \hat{a}_{b-} \tag{3.4}$$

where b is the bandwidth, $K_b(x) = K(x/b)/b$ and $K()$ is a kernel function, and $\mathit{1}()$ denotes an indicator function.

Samples and implementation. Our main analysis of the policy effect of increased ethanol use on ozone concentrations restricts hourly observations to afternoon readings between 12pm and 4pm, when the rate of ozone production is at its highest (Sections 2 and 3). For the other pollutants, we consider observations during the evening commuting hours from 5pm to 8pm. In short, we look for effects from the energy policy where we are more likely to find them. In sensitivity analysis, we examine other hour windows for the different pollutants.

In the first step of estimation, we restrict observations to plus and minus 90 days of the cutoff point (for the given discontinuity), namely three days after the mandated deadline for distributor shipments, to allow fuel stations and consumers to adjust, as explained.[33] In the second step, we use CCT's bandwidth selection criterion to determine the number of days before and after the discontinuity. The tested bandwidth turns out to be about 30 days, ensuring that the 90 days we consider in the first step suffices for the second step. Of note, the first and second discontinuities occurred three months apart. As we examine each discontinuity one by one, we do not vary the 180-day window, with the objective of making our estimates comparable across discontinuities.

3.5 Results

As discussed in Section 3, we face the empirical challenge that, although gasoline fuel accounted for the majority share of Sao Paulo's 6 million cars and 1 million motorcycles during each of the four discontinuities, changes to the composition of gasoline were not large. Also, ozone's reactivity and sensitivity

[33]See Table H.1. As a robustness check, we allow the mandated ethanol shift to phase in over several days, rather than abruptly. As explained, a fuzzy design would require that we observe the exact composition of downstream fuel inventories—the "compliance rate"—over these few days.

to the immediate surroundings implies that concentrations can be quite variable for unobserved, idiosyncratic reasons. This is an empirical challenge that Auffhammer and Kellogg (2011) likely also faced. In our setting, this variability can immediately be seen from the scatter in Figure I.4 showing demeaned log afternoon ozone concentrations at the level of the data—a site-hour-date triple—following the first step of the local linear regression estimation procedure (Section 4). The dots indicate log concentration residuals, $\tilde{y}$, with all covariates accounted for, including meteorology and the eight-order polynomial trend over the whole 180-day period, flexibly by site-hour. To emphasize, we estimate models separately by discontinuity throughout the analysis, and "pool" across discontinuities only in a robustness test. The fitted lines on either side of each discontinuity are only for illustration, as here they are based on a fixed bandwidth of 30 days rather than the bandwidth tests reported below.

Before presenting results for our preferred local linear regression specification, Table H.3 presents estimates for regression model (1), estimated from the 180-day sample in a single step by OLS with $treat_t$ included. The estimated effect on ambient ozone concentrations from an increase in the ethanol content in gasoline is positive and significant at least at the 5% level across all four discontinuities. The average point estimate and standard error across the four discontinuities is 0.125 and 0.042 log points, respectively. (Standard errors shown in Table H.3 are one-way clustered at the site-date level.) Estimates of the increase in afternoon ozone levels—a 0.11 to 0.16 increase in log points—caused by a 5 percentage point increase in the ethanol fraction appear quite large (see below). Given concerns about the extent to which high-order polynomials can single-handedly account for unobserved confounding factors, such as seasonality, as we move away from the cutoff, we follow the advice of Gelman and Imbens (2014) and turn to local linear regression.

Consider again demeaned log ozone concentrations, demeaned separately by 180-day sample, at the site-hour-date level, $\tilde{y}_{iht}$. To account for correlation across hourly readings within date and site (i.e., to "cluster" our inference at the site-date level), we then take the mean of the demeaned log concentrations across hours within site-date. It is on these residual log concentrations at the site-date level that we implement local linear regression. (We later check robustness to implementing local linear regression directly on $\tilde{y}_{iht}$.) Table H.4, panel A presents local linear regression estimates for the effect on mean afternoon ozone concentrations from raising the ethanol content in gasoline fuel, under alternative bandwidths according to either the CCT or the IK criteria (Section 4). Selected bandwidths range between 22 and 35 days and turn out to be similar across the two criteria, for example, 22 (CCT) and 24

(IK) days on either side of the discontinuity # 3 cutoff. Importantly, a 30-day or so bandwidth coupled with 12 sites (with sites added in 2012) implies that the number of observations is not large—between 500 and 1,000.

Across all four discontinuities and both bandwidth selection criteria, the local regression's estimated effect of ethanol on mean afternoon ozone levels is positive and somewhat lower than direct estimates of (1) using the whole 180-day period (Table H.3). Point estimates range from 0.041 (discontinuity # 4, IK) to 0.126 (discontinuity # 3, CCT), and average 0.090 log points across discontinuities and bandwidth criteria (with an average standard error of 0.035 log points). Point estimates under the IK criterion average 0.083 compared with 0.097 log points under CCT. These estimated magnitudes, based on a 5 percent ethanol increase in fuel content among 70% of light vehicles, that tended to be burning blended gasoline at the cutoff dates (Table H.1), are comparable to—if slightly higher than—point estimates reported by Salvo and Geiger (2014, Figure 4) for a shift from E25 to E100 among 60% of bi-fuel vehicles (bi-fuel vehicles accounting for 50% of VMT by passenger cars).

Robustness. In Table H.4, panel B we alternatively take the maximum of each afternoon's (log) ozone concentrations as the dependent variable, with the resulting data at the site-date level, and repeat the estimation routine, namely demeaning the dependent variable using a 180-day sample followed by implementing local linear regression on a sample of bandwidth tested according to each of the two criteria. The estimated effect of ethanol on maximum ozone levels is again positive across all four discontinuities and both bandwidth criteria, averaging 0.081 log points.

Figure I.5 shows that the estimated positive effect of ethanol on mean afternoon ozone concentrations, reported in Table H.4A is not overly sensitive to varying the bandwidth of days around the cutoff point. 95% confidence intervals are quite stable and remain above (or almost above) zero as we decrease or increase the bandwidth starting at the level selected by either the CCT or the IK criteria. Perhaps unsurprisingly, estimates can become rather less stable, with confidence bands widening and shifting, as we reduce the bandwidth by 10 days. Estimates are somewhat more precise as we increase the bandwidth, though this comes at the potential expense of introducing omitted variable bias.

Figure I.6 tests robustness to the demeaning of the data (at varying bandwidths). Solid lines indicate point estimates for the treatment effect following the procedure described above (demean the dependent variable at the observed level of the data, a site-hour-date triple, and take the mean of the demeaned values across hours within site-date). Dashed lines indicate point estimates under an alternative procedure, whereby we implement local linear regression directly on site-hour-date level demeaned values, rather than on site-date level means of these values.[34]

Our findings are further robust to pooling the demeaned data (that goes into the second step) across the four discontinuities, while still demeaning separately by 180-day sample in the first step to flexibly allow for, say, seasonal variation: Table H.5 reports an estimated effect of 0.081 to 0.083 log points.

Table H.6 reports robustness to hour-of-the-day—mean over morning hours from 7am to 11am, or the maximum 8-hour average (e.g., Lin et al. 2001)—and to functional form—ozone, not log ozone, concentration. In panel D, and pooling ozone residuals across the four 180-day samples, the effect of raising the ethanol fraction on the daily maximum 8-hour ozone average is 4.8 μg/m^3 (s.e. 1.2 μg/m^3), or 8.0% of the sample mean of daily maximum 8-hour averages, of 59.9 μg/m^3.

We also implement local linear regression separately by ozone monitoring site (and discontinuity). Each panel of Figure I.7 plots the distribution of 53 (12 + 12 + 12 + 17) point estimates across sites and discontinuities. Most density lies on the positive domain, with medians of 0.084 and 0.063 log points,

[34]Dotted lines in the figure indicate point estimates for the treatment effect under yet another procedure, whereby we first take means of the data across the afternoon hours from 12pm to 4pm, and only then proceed to demean the site-date varying mean afternoon log ozone concentrations.

respectively under the CCT and IK criteria (a median of 0.072 log points overall).

Finally, Appendix Table H.8 reports robustness to varying the order of the polynomial trends specified over each 180-day first-step sample. Appendix Table H.9 reports bootstrap standard errors, in which a bootstrap sample is a sample of site-date pairs, to account for sampling variation in the first step. As an alternative to clustering at the site-date level, we alternatively bootstrap the sample at the site-week level. In the second step, we alternatively fix the bandwidth selected from the original sample, or select the bandwidth according to each bootstrap sample. Appendix Table H.11 considers specifications in which we allow the mandated ethanol shift to phase in over a few days, rather than abruptly.

Other robustness tests—not shown for brevity—include widening the first-step sample beyond 180 days, or specifying a uniform rather than triangular kernel. We add variables (sunshine hours) or high-order terms (quartic rather than quadratic) in the vector of meteorological controls, or control for meteorology in levels (not log) or bins (temperature every 1^0C, radiation every 150 W/m^2). We include, as separate controls, monthly gasoline and ethanol fuel shipments reported by distributors to the state market—while too aggregate in time and space, these are the quantity data that are available (Figure I.1(c)).

Our findings are also robust to specifying the cutoff point at four days (rather than our preferred three) from the mandated distributor deadline.

Ethanol and PM2.5, CO and NOx. Table H.7 provides local linear regression estimates for the effect on ambient PM2.5, CO and NOx levels from raising gasoline fuel's ethanol content—see panels A, B and C, respectively. We follow exactly the same procedure as we implemented when evaluating the policy effect on ozone (Table H.4A), recalling that now we consider evening commuting hours of peak traffic congestion, between 5pm and 8pm (Table H.2), rather than afternoon hours of peak ozone formation.

Averaging across discontinuities and bandwidth criteria, point estimates in log points are -0.034 for PM2.5, -0.019 for CO and -0.059 for NOx. Estimates tend to be smaller in magnitude than in the case for ozone, and they tend not to be significantly different than zero. Estimates on NOx are less precise than in the case for ozone, i.e., standard errors across discontinuities and bandwidth criteria average 0.057 and 0.035 log points for NOx and ozone, respectively, with inference in both cases relying on a similar number of monitors in the metropolis. Pooling residuals across all 180-day samples in a single (second-step) regression, point estimates for NOx are a statistically negative -0.077 log points under the CCT criterion. Salvo and Geiger (2014) found NO levels to fall with increased ethanol penetration and their estimates for NO were less precise compared to ozone.

PM2.5 levels do *not* appear to fall with higher ethanol penetration, or at least we are unable to detect a fall, as our inference relies on a low number of monitors (one and three for the last two discontinuities). Estimates for other hours (morning commuting hours, 7am to 11am) and separate NO and NO_2 components similarly tend to be smaller in magnitude and less precise than for ozone (not shown for brevity). Unfortunately, atmospheric levels of VOCs such as formaldehyde, acetaldehyde and ethanol are not routinely monitored (Nogueira et al. 2014).

3.6 Conclusion

This paper exploits four sharp discontinuities in the ethanol component of blended gasoline sold across Brazil to assess the effect on air quality in the Sao Paulo metropolis from raising the fraction of ethanol in the gasoline-ethanol light transportation fuel mix. These ethanol policy changes, each amounting to 5 percentage points by volume, were spaced at least three months apart and were mandated by the central government not as a response to urban air pollution, but to intervene in the quantity and price outcomes of the sugar industry (sugar and ethanol products to domestic and export markets), or simply to reverse previous changes. We test that the blend discontinuities did not lead to price discontinuities, and argue that composition changes were not salient to fuel consumers (and, if they were, likely responses would help us detect the effect of increased ethanol on urban air outcomes). Using available data, we check that on each occasion, blended gasoline accounted for the majority of distance traveled by the fleet of passenger cars and motorcycles in the state of Sao Paulo. Our focus is on ozone, which often exceeds the local 8-hour air quality standard (71 ppb, comparable to the US EPA's 70 ppb), is the one regulated pollutant whose levels have not trended downward in the past decade, and whose relevance is likely to grow in a warming climate. We also test for effects on other monitored species, PM2.5 in particular.

Across the four discontinuities, which we examine separately to allow for unobserved heterogeneity, we find that ozone levels rose by about 8 percent as the gasoline blend shifted from E20 to E25, i.e., toward a higher ethanol fraction. Such observational evidence (1) is consistent (in sign and magnitude) with the results of a recent empirical study that exploited consumers smoothly substituting between E20/E25 gasoline and E100 ethanol, (2) is theoretically consistent with ozone formation over urban Sao Paulo being VOC-limited, (3) enables the credible inference of the health damage of ozone that, at 70 ppb, is hotly debated, and (4) is opposite in sign to the predictions of an atmospheric modeling study for Sao Paulo, based on emissions inventories and modeled chemical and physical reactions, that suggested sizable ozone abatement from a *hypothetical* shift to ethanol by the light-vehicle fleet. In contrast to ozone, point estimates of changes to PM2.5 are small in magnitude but confidence intervals are wide.

Since the ethanol-blended gasoline policy changes were common across the country, future research can examine their effect on ozone pollution at locations outside of the Sao Paulo metropolis. For example, over the period 2010 to 2013, ozone monitors were in place in about 15 cities in the state of Sao Paulo, with distances ranging from 100 to 600 km from the metropolis, and populations varying by almost an order of magnitude (0.1 to 0.6 million). To the extent that atmospheric ozone production regimes differ by locality (and possibly season)—to be proxied, for example, by weekend-weekday ozone concentration ratios, population size, and/or prevailing winds from the Sao Paulo megacity transporting NOx-rich air parcels over varying distances—one may examine heterogeneous effects of increased ethanol penetration on ambient ozone levels.[35]

In a country where sugarcane ethanol is commonly perceived by consumers and described by policymakers as a policy to abate gasoline emissions and improve air quality (e.g., Salvo and Huse 2013, CBN Noticias 2010, Angelo 2012),

[35] Preliminary analysis suggests that this research direction is promising. We are grateful to a Reviewer for suggesting it.

we hope that our research will better inform the public discourse. Our research is also supportive of recent calls for a review of current vehicle emissions standards that allow the deduction of the mass of unburned ethanol, itself a reactive VOC, emitted from the tailpipe. We caution that our findings do not speak to other potential life-cycle benefits and costs of substituting ethanol (current or future generation) for gasoline, nor do they make assumptions on how on-board emissions control technologies might evolve. Moreover, we caution that our analysis applies only to pollutants that are currently routinely monitored by Sao Paulo's (and other) environmental authorities, thus ignoring possible effects, in particular, on particle levels in the submicron ($< 1\mu$m or PM1) and nanometer ($< 0.1\mu$m or PM0.1) size ranges.

Beyond Sao Paulo, our findings may be relevant to urban areas with a VOC-limited atmosphere—which the literature suggests is common to many densely populated cities in the US, Europe and Asia—and countries that are planning to increase ethanol blending requirements, whether to meet local regulation or international agreements, such as the RFS in the US and the INDCs pledged during COP21 in Paris.

Bibliography

Aizer, Anna, and Flvio Cunha. (2012). The Production of Human Capital: Endowments, Investments and Fertility. National Bureau of Economic Research Working Paper No. 18429.

Anderson, L. G. (2009). Ethanol fuel use in Brazil: Air quality impacts. *Energy & Environmental Science*, 2, pp. 1015-1037.

Angelo, C. (2012). News in Focus: Growth of ethanol fuel stalls in Brazil. *Nature*, 491, p. 646.

Arceo, E., Hanna, R., Oliva, P. (2016). Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. *Economic Journal*, 126(591), pp. 257-280.

Attanasio, O., Lechene, V. (2002). Tests of income pooling in household decisions. *Review of economic dynamics*, 5(4), pp. 720-748.

Auffhammer, M., Kellogg, R. (2011). Clearing the air? The effects of gasoline content regulation on air quality. *American Economic Review*, 101(6), pp. 2687-2722.

Balakrishnan, T.R., K. Vaninadha Rao, Evelyne Lapierre-Adamcyk, and Karol J. Krotki (1987). A Hazard Model Analysis of the Covariates of Marriage Dissolution in Canada. *Demography*, 24(3), pp. 395-406

Banks, James, Richard Blundell, and Arthur Lewbel (1997). Quadratic Engel curves and consumer demand. *Review of Economics and Statistics*, 79(4), pp. 527-539.

Basu, K. (2006). Gender and say: A model of household behaviour with endogenously determined balance of power. *The Economic Journal*, 116(511), pp. 558-580.

Becker, Gary. (1974). A Theory of Social Interactions. *Journal of Political Economy*, 82(6), pp. 106394

Becker, Gary. (1981). *A Treatise on the Family.* Cambridge: Harvard University Press, 1981, enlarged edition 1991.

Bennett, Neil G., Ann K. Blanc, and David E. Bloom. (1988). Commitment and the Modern Union: Assessing the Link between Premarital Cohabitation and Subsequent Marital Stability. *American Sociological Review*, 53(1), pp. 127-138

Blundell, Richard, and Jean Marc Robin (1999). Estimation in large and disaggregated demand systems: an estimator for conditionally linear systems. *Journal of Applied Econometrics*, 14(3), pp. 209-232.

Blundell, R., P.A. Chiappori, T. Magnac, and C. Meghir. (2007). Collective labor supply: Heterogeneity and non-participation. *Review of Economic Studies*, 74, pp. 417-445.

Blundell, Richard, Pierre André Chiappori, and Costas Meghir (2005). Collective labor supply with children. *Journal of political Economy*, 113(6), pp. 1277-1306.

Bobonis, G. (2009). Is the allocation of resources within the household efficient? New evidence from a randomized experiment. *Journal of Political Economy*, 117(3), 453-503.

Bourguignon, Franois, Martin Browning, and Pierre-André Chiappori (2009). Efficient intra-household allocations and distribution factors: implications and identification. *The Review of Economic Studies*, 76(2), pp. 503-528.

Bracher, Michael, Gigi Santow, S. Phillip Morgan, and James Trussell. 1993.

Marriage Dissolution in Australia: Models and Explanations. *Population Studies*, 47, pp.403-425

Brien, Michael J., Lee A. Lillard, and Steven Stern (2006). Cohabitation, Marriage, and Divorce in a Model of Match Quality. *International Economic Review*, 47(2), pp. 451-94.

Brown, Susan, Wendy D. Manning, and Krista K. Payne (2015). Relationship quality among cohabiting versus married couples. *Journal of Family Issues*, 0192513X15622236, pp. 1-24.

Browning, Martin, et al (1994). Income and outcomes: A structural model of intrahousehold allocation. *Journal of Political Economy*, pp. 1067-1096.

Browning, Martin, and Pierre-André Chiappori (1998). Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica*, pp. 1241-1278.

Calonico, S., Cattaneo, M. D., Titiunik, R., (2014a). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 14(4), pp. 909-946.

Calonico, S., Cattaneo, M. D., Titiunik, R., (2014b). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6), pp. 2295-2326.

CBN Noticias (2010). *Interview with Dr Jose Goldemberg, former Secretary of the Environment.* Radio interview aired on January 5, 2010, Sao Paulo.

CETESB (2011, 2012, 2013). *Relatorio Anual sobre a Qualidade do Ar no Estado de Sao Paulo.* In Portuguese: Annual Report on Air Quality in the State of Sao Paulo. Companhia Ambiental do Estado de Sao Paulo (Sao Paulo State Environmental Protection Agency), Sao Paulo.

Chen, Y., Ebenstein, A., Greenstone, M., Li, H., (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. *Proceedings of the National Academy of Sciences*, 110(32), pp. 12936-12941.

Cherlin, A. J. (2004). The deinstitutionalization of American marriage. *Journal of Marriage and Family*, 66(4), pp. 848-861.

Cherlin, Andrew. (2009). The Marriage-Go-Round: The State of Marriage and the Family in America Today. New York: Alfred A. Knopf.

Chiappori, Pierre-André. (1988). Rational Household Labor Supply. *Econometrica*, 56(1), pp. 63-89.

Chiappori, Pierre-André. (1992) Collective Labor Supply and Welfare. *Journal of Political Economy*, 100(3), pp. 437-467.

Chiappori, Pierre-André, Bernard Fortin, and Guy Lacroix (2002). Marriage market, divorce legislation, and household labor supply. *Journal of political Economy*, 110(1) pp. 37-72.

Chicago Tribune (2015). Ald. Burke's lobbying bonanza ethanol bill is spared under council 'kill ordinance'. Reported by John Byrne, May 20.

Cook, R., Phillips, S., Houyoux, M., Dolwick, P., Mason, R., Yanca, C., Zawacki, M., Davidson, K., Michaels, H., Harvey, C., Somers, J., Luecken, D. (2011). Air quality impacts of increased use of ethanol under the United States' Energy Independence and Security Act. *Atmospheric Environment*, 45(40), pp. 7714-7724.

Cunha, Flavio, and James Heckman. (2007). The Technology of Skill Formation, *American Economic Review*, 97(2), pp. 31-47.

Anyck Dauphin, Abdel-Rahmen El Lahga, Bernard Fortin, Guy Lacroix. (2011). Are children decision makers within the household? *The Economic Journal*, 121(553), pp. 871-903.

Davis, L. W., (2008). The effect of driving restrictions on air quality in Mexico City. *Journal of Political Economy*, 116(1), pp. 38-81.

de Gouw, J. A., Gilman, J. B., Borbon, A., Warneke, C., Kuster, W. C., Goldan, P. D., Holloway, J. S., Peischl, J., Ryerson, T. B., Parrish, D. D., Gentner, D. R., Goldstein, A. H., Harley, R. A. (2012). Increasing atmospheric burden of ethanol in the United States. *Geophysical Research Letters*, 39(15), L15803.

Deschenes, O., Greenstone, M., Shapiro, J. S. (2012). Defensive investments and the demand for air quality: Evidence from the NOx Budget Program and ozone reductions. National Bureau of Economic Research, Working Paper No. 18267.

De Vos, S. (1999). Comment of coding marital status in Latin America. *Journal of Comparative Family Studies*, 30, pp. 79-93.

Dominici, F., Greenstone, M., Sunstein, C. R. (2014). Particulate Matter matters. *Science*, 344(6181), pp. 257-259.

Durbin, T. D., Miller, J. W., Younglove, T., Huai, T., Cocker, K. (2007). Effects of fuel ethanol content and volatility on regulated and unregulated exhaust emissions for the latest technology gasoline vehicles. *Environmental Science & Technology*, 41, pp. 4059-4064.

Eizenberg, A., Salvo, A. (2015). The rise of fringe competitors in the wake of an emerging middle class: An empirical analysis. *American Economic Journal: Applied Economics*, 7(3), pp. 85-122.

Finlayson-Pitts, B., Pitts, J. (2000). *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications*. Academic Press, San Diego.

Fujita, E. M., Campbell, D. E., Stockwell, W. R., Saunders, E., Fitzgerald, R., Perea R. (2016). Projected ozone trends and changes in the ozone precursor relationship in the South Coast Air Basin in response to varying reductions of precursor emissions. *Journal of the Air & Waste Management Association*, 66(2), pp. 201-214.

Furstenberg, Frank F., Jr. (1996). The future of marriage. *American Demographics*, 18(6), pp. 34-40.

Gelman, A., Imbens, G., (2014). Why high-order polynomials should not be used in regression discontinuity designs. National Bureau of Economic Research, Working Paper No. 20405.

Gemici, Ahu, and Steve Laufer. (2011). Marriage and cohabitation. New York University, mimeo (2011).

Ginnebaugh, D. L., Liang, J., Jacobson, M. Z. (2010). Examining the temperature dependence of ethanol (E85) versus gasoline emissions on air pollution with a largely-explicit chemical mechanism. *Atmospheric Environment*, 44(9), pp. 1192-1199.

Hahn, J., Todd, P., Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1), pp. 201-209.

He, J., Gouveia, N., Salvo, A. (2016). The unintended consequences of policies to abate truck flows inside the Sao Paulo megacity. Mimeo, National University of Singapore and University of Sao Paulo.

Heuveline, P. and J.M. Timberlake. (2004). The Role of Cohabitation in Family Formation: The United States in Comparative Perspective. *Journal of Marriage and Family*, 66(5) pp. 1214-1230.

Hochhauser, A. M., Schleyer, C. H. (2014). Summary of Research on the Use of Intermediate Ethanol Blends in On-Road Vehicles. *Energy & Fuels*, 28(5), pp. 3236-3247.

Hoem, Britta and Jan Hoem. (1992). The Disruption of Marital and Nonmarital Unions Contemporary Sweden. In *Demographic Applications of Event*

History Analysis, James Trussell, Richard Hankinson, and Judith Tilton, 61-93. Oxford England: Clarendon Press.

Hubbard, C. P., Anderson, J. E., Wallington, T. J. (2014). Ethanol and air quality: influence of fuel ethanol content on emissions and fuel economy of flexible fuel vehicles. *Environmental Science & Technology*, 48(1), pp. 861-867.

Imbens, G. W., Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies* 79, pp. 933-959.

Imbens, G. W., Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2), pp. 615-635.

Jacob, D. J. (1999). *Introduction to Atmospheric Chemistry.* Princeton University Press, Princeton.

Jacobson, M. Z. (2007). Effects of ethanol (E85) versus gasoline vehicles on cancer and mortality in the United States. *Environmental Science & Technology*, 41(11), pp. 4150-4157.

Jing, P., Lu, Z., Xing, J., Streets, D., Tan, Q., O'Brien, T., Kamberos, J. (2014). Response of the summertime ground-level ozone trend in the Chicago area to emission controls and temperature changes, 2005-2013. *Atmospheric Environment*, 99, pp. 630-640.

Lee, Jungmin. Marriage, the sharing rule, and pocket money: The case of South Korea. *Economic Development and Cultural Change*, 55(3), pp. 557-581.

Lichter, Daniel T., and Zhenchao Qian. (2008). Serial Cohabitation and the Marital Life Course. *Journal of Marriage and Family*, 70, pp. 861-878.

Lillard, Lee A., Michael J. Brien, and Linda J. Waite. (1995). Premarital Cohabitation and Subsequent Marital Dissolution: A Matter of Self-Selection? *Demography*, 32(3), pp. 437-457.

Kiernan, K. E. (2000) European Perspectives on Union Formation. in Waite, L., Bachrach, C., Hindin, M., Thomson, E., and Thornton, A., (eds) *Ties that Bind: Perspectives on Marriage and Cohabitation.* Hawthorne: Aldine de Gruyter.

Lin, C.-Y. C., Jacob, D. J., Munger, J. W., Fiore, A. M. (2000). Increasing background ozone in surface air over the United States. *Geophysical Research Letters*, 27(21), pp. 3465-3468.

Lin, C.-Y. C., Jacob, D. J., Fiore, A. M. (2001). Trends in exceedances of the ozone air quality standard in the continental United States, 1980-1998. *Atmospheric Environment*, 35, pp. 3217-3228.

Lin, C.-Y. C. (2010). A spatial econometric approach to measuring pollution externalities: An application to ozone smog. *Journal of Regional Analysis and Policy*, 40(1), pp. 1-19.

Lin Lawell, C.-Y. C., Zhang, W., Umanskaya, V. I. (2016). The effects of driving restrictions on air quality: Theory and empirical evidence. Mimeo, University of California at Davis.

Lixing Li, and Xiaoyu Wu (2011). Gender of children, bargaining power, and intrahousehold resource allocation in China. *Journal of Human Resources*, 46(2), pp. 295-316.

Lofquist, D., Lugaila, T., O'Connell, M., Feliz, S (2012). Households and families: 2010: 2010 Census briefs (C2010BR-14). *Retrieved from http://www.census.gov/prod/cen2010/briefs/c2010br-14.pdf*

Ludwig, J., Miller, D. L. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics*, 122(1), pp. 159-208.

Lundberg, Shelly, and Robert A. Pollak. (1994). Noncooperative bargaining models of marriage. *The American Economic Review*, pp. 132-137.

Lundberg, Shelly J., Robert A. Pollak, and Terence J. Wales (1997). Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit. *Journal of human resources*, pp. 463-480.

Lundberg, Shelly J., Robert A. Pollak. (2008). Family Decision-Making. *The New Palgrave Dictionary of Economics*, 2nd edition

Lundberg, Shelly, and Robert A. Pollak. (2013). Cohabitation and the Uneven Retreat from Marriage in the US, 1950-2010. *Human Capital in History: The American Record*. University of Chicago Press

Madronich, S. (2014). Atmospheric chemistry: Ethanol and ozone. *Nature Geoscience*, 7, pp. 395-397.

Martins, L. D., Andrade, M. F. (2008a). Emission scenario assessment of gasohol reformulation proposals and ethanol use in the metropolitan area of Sao Paulo. *The Open Atmospheric Science Journal*, 2, pp. 166-175.

Martins, L. D., Andrade, M. F. (2008b). Ozone formation potentials of volatile organic compounds and ozone sensitivity to their emission in the megacity of Sao Paulo. *Water, Air, & Soil Pollution*, 195, pp. 201-213.

Martins, L. D., Andrade, M. F., Ynoue,R. Y. (2008). Ambient volatile organic compounds in the megacity of Sao Paulo. *Quim. Nova*, 31(8), pp. 2009-2013.

Matouschek, N., Rasul, I. (2008). The economics of the marriage contract: Theories and evidence. *Journal of Law and Economics*, 51(1), pp. 59-110.

Mazzocco, Maurizio (2007). Household intertemporal behaviour: A collective characterization and a test of commitment. *The Review of Economic Studies*, 74(3), pp. 857-895.

McElroy, Marjorie B. and Mary Jean Horney. (1981) Nash Bargained Household Decisions. *International Economic Review*, 22(2), pp. 333 349

Negri, J. C. (2010). *Balanco Energetico do Estado de Sao Paulo*. In Portuguese: Energy Balance for the State of Sao Paulo. Secretaria de Saneamento e Energia do Governo de Sao Paulo (Sao Paulo State Government, Secretariat for Sanitation and Energy), Sao Paulo.

Nogueira, T., Dominutti, P. A., Carvalho, L. R. F., Fornaro, A., Andrade, M. F. (2014). Formaldehyde and acetaldehyde measurements in urban atmosphere impacted by the use of ethanol biofuel: Metropolitan Area of Sao Paulo (MASP), 2012-2013. *Fuel*, 134, pp. 505-513.

Nogueira, T., Souza, K. F., Fornaro, A., Andrade, M. F., Rothschild, L., Carvalho, L. R. F. (2015). On-road emissions of carbonyls from vehicles powered by biofuel blends in traffic tunnels in the Metropolitan Area of Sao Paulo, Brazil. *Atmospheric Environment*, 108, pp. 88-97.

Nopmongcol, U., Griffin, W. M., Yarwood, G., Dunker, A. M., MacLean, H. L., Mansell, G., Grant, J. (2011). Impact of dedicated E85 vehicle use on ozone and particulate matter in the US. *Atmospheric Environment*, 45(39), pp. 7330-7340.

Olley, G. S., and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64, pp. 1263-1297

Orlando, J. P., Alvim, D. S., Yamazaki, A., Correa, S. M. & Gatti, L. V. (2010). Ozone precursors for the Sao Paulo metropolitan area. *Science of the Total Environment*, 408, pp. 1612-1620.

Parrado, Emilio A. and Marta Tienda. (1997). Womens Roles and Family Formation in Venezuela: New Forms of Consensual Unions? *Social Biology*, 44, pp. 1-24.

Pereira, P. A. P., Santos, L. M. B., Sousa, E. T., Andrade, J. B. (2004). Alcohol- and gasohol-fuels: a comparative chamber study of photochemical ozone formation. *Journal of the Brazilian Chemical Society*, 15, pp. 646-651.

Ramey, Garey and Valerie A. Ramey. (2010). The Rug Rat Race. *Brookings Papers on Economic Activity*, Spring, pp. 129-176.

Raymo, J. M., Iwasawa, M., Bumpass, L. (2009). Cohabitation and Family Formation in Japan. *Demography*, 46(4), pp. 785803.

RFA (2016). *US Ethanol Industry Outlook.* Renewable Fuels Association, Washington, DC.

Salvo, A. (2016). Flexible fuel vehicles, less flexible minded consumers: Price salience experiments at the pump. Mimeo, National University of Singapore.

Salvo, A., Huse, C. (2011). Is arbitrage tying the price of ethanol to that of gasoline? Evidence from the uptake of flexible-fuel technology. *Energy Journal*, 32(3), pp. 119-148.

Salvo, A., Huse, C. (2013). Build it, but will they come? Evidence from consumer choice between gasoline and sugarcane ethanol. *Journal of Environmental Economics and Management*, 66, pp. 251-279.

Salvo, A., Geiger, F. M. (2014). Reduction in local ozone levels in urban Sao Paulo due to a shift from ethanol to gasoline use. *Nature Geoscience*, 7, pp. 450-458.

Samuelson, Paul A. (1956). Social Indifference Curves. *Quarterly Journal of Economics*, 70(1), pp. 122.

Seinfeld, J. H., Pandis, S. N. (1998). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. Wiley, New York.

Shonkwiler, J. Scott, and Steven T. Yen (1999). Two-step estimation of a censored system of equations. *American Journal of Agricultural Economics*, 81(4), pp. 972-982.

Stedman, D. H. (2004). Photochemical ozone formation, simplified. *Environmental Chemistry*, 1, pp. 65-66.

Stephens, S., Madronich, S., Wu, F., Olson, J. B., Ramos, R., Retama, A., Munoz, R. (2008). Weekly patterns of Mexico City's surface concentrations of CO, NOx, PM10 and O3 during 19862007. *Atmospheric Chemistry and Physics*, 8, pp. 5313-5325.

Suarez-Bertoa, R., Zardini, A. A., Platt, S. M., Hellebust, S., Pieber, S. M., El Haddad, I., Temime-Roussel, B., Baltensperger, U., Marchand, N., Prevot, A. S. H., Astorga, C. (2015). Primary emissions and secondary organic aerosol formation from the exhaust of a flex-fuel (ethanol) vehicle. *Atmospheric Environment*, 117, pp. 200-211.

Thomas, Duncan, Dante Contreras, and Elizabeth Frankenberg (2002). Distribution of power within the household and child health. Santa Monica: RAND. Mimeo.

Thornton, Arland, and Linda Young-DeMarco (2001). Four Decades of Trends in Attitudes Toward Family Issues in the United States: The 1960s through the 1990s. *Journal of Marriage and Family*, 63(4), pp. 1009-1037.

Todd, Petra E., and Kenneth I. Wolpin. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement, *Economic Journal*, 113(485), pp. F3 - F33.

Todd, Petra E., and Kenneth I. Wolpin. (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps, *Journal of Human Capital*, 1(1), pp. 91-136.

Udry, Christopher (1996). Gender, agricultural production, and the theory of the household. *Journal of Political Economy*, pp. 1010-1046.

US EPA (2004). *Evaluating Ozone Control Programs in the Eastern United States*. United States Environmental Protection Agency, Office of Atmospheric Programs, Washington, DC.

Wallington, T. J., Anderson, J. E., Kurtz, E. M., Tennison, P. J. (2016).

Biofuels, vehicle emissions, and urban air quality. *Faraday Discussions*, 189, pp. 121-136.

Weiss, Yoram. (1997). The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And What Happens Upon Divorce. in Mark R. Rosenzweig and Oded Stark, eds. *Handbook of Population and Family Economics*, 1, 81-123.

Weiss, Yoram, and Robert J. Willis (1997). Match quality, new information, and marital dissolution. *Journal of Labor Economics*, 15(1), pp. S293-S329.

Van de Kaa, D. J. (1987). Europes Second Demographic Transition. *Population Bulletin*, 42 (1), Washington, The Population Reference Bureau.

Voena, Alessandra. (2015). Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples? *American Economic Review*, 105(8), pp. 2295-2332.

Appendix A

Chapter 1: Tables

Table A.1: Summary statistics

	(1) Husband	(2) Wife	(3) Household
Number of observations			10368
Have one child in house			49.5%
Average age	41.1	37.4	
Labor supply rate	92.8%	67.0%	
Car ownership rate			35.5%
House ownership rate			60.2%
Living in rural area			23.9%
Living in North area			56.7%
Budget share for clothes			13.8%
Children's clothing share			34.6%
Female clothing share			
-with one child			55.1%
-with no children			54.1%

Notes: The sample consists of all households with one head, one spouse and at most one child living in house. Column (1) and column (2) are based on individuals. Column (3) is based on households. Budget share for clothes is defined as yearly expenditure on clothes divided by yearly income. Children's clothing share is defined as children's clothing expenditure divided by total clothing expenditure within a household. Female clothing share is defined as female clothing expenditure divided by summation of female clothing expenditure and male clothing expenditure.

Table A.2: Effects of female income share on clothing expenditure share

	(1) with child	(2) without child
Panel A: Without control function		
children's clothing budget share	**0.080	
	(0.036)	
female clothing budget share	*0.086	*0.084
	(0.052)	(0.050)
Panel B: With control function		
children's clothing budget share	0.054	
	(0.099)	
female clothing budget share	0.036	0.079
	(0.109)	(0.105)

Notes: *** p<0.01, ** p<0.05, * p<0.1. The results show effects of female income share on clothing expenditure share. Panel A shows results without controlling individual unobserved characteristics. Panel B shows results control for individual unobserved characteristics. Column (1) shows results for households with one child. Column (2) shows results with no children. Distributor factors include female income share, age and education difference of husband and wife, and head sex dummy. Demographic information entering the utility function include monthly household total income, female age and education, children's age and sex if applicable, labor supply decision, car ownership, house ownership, household points, clothing price, rural fixed effect, monthly fixed effect and state fixed effect.

Table A.3: Effects of female income share on clothing expenditure share by wealth level

	(1)	(2)	(3)	(4)
	with children		without children	
	rich	poor	rich	poor
Panel A: Without control function				
children's clothing budget share	0.042	**0.094		
	(0.046)	(0.039)		
female clothing budget share	0.051	*0.100	0.071	*0.088
	(0.068)	(0.057)	(0.070)	(0.053)
Panel B: With control function				
children's clothing budget share	0.024	0.066		
	(0.103)	(0.102)		
female clothing budget share	-0.007	0.047	0.061	0.082
	(0.117)	(0.111)	(0.123)	(0.105)

Notes: *** p<0.01, ** p<0.05, * p<0.1. The results show effect of female income share on clothing expenditure share. Panel A shows results without controlling individual unobserved characteristics. Panel B shows results control for individual unobserved characteristics. Column (1) shows results for rich households with one child. Column (2) shows results for poor households with one child. Column (3) shows results for rich households with no children. Column (4) shows results for poor households with no children. Households are categorized as rich or poor based on household points. Distributor factors include female income share, age and education difference of husband and wife, and head sex dummy. Demographic information entering utility function include monthly household total income, female age and education, children's age and sex if applicable, labor supply decision, car ownership, house ownership, household points, clothes price, rural fixed effect, monthly fixed effect and state fixed effect.

Table A.4: Effect of female income share on clothing expenditure by rural and urban areas

	(1)	(2)	(3)	(4)
	with children		without children	
	Urban	Rural	Urban	Rural
Panel A: Without control function				
children's clothing budget share	** 0.093	0.038		
	(0.039)	(0.056)		
female clothing budget share	*0.111	0.004	**0.108	0.014
	(0.057)	(0.084)	(0.053)	(0.079)
Panel B: With control function				
children's clothing budget share	0.075	0.006		
	(0.101)	(0.111)		
female clothing budget share	0.073	-0.057	0.101	-0.010
	(0.112)	(0.124)	(0.106)	(0.124)

Notes: *** p<0.01, ** p<0.05, * p<0.1. The results show effect of female income share on clothing expenditure share. Panel A shows results without controlling individual unobserved characteristics. Panel B shows results control for individual unobserved characteristics. Column (1) shows results for households with one child in urban area. Column (2) shows results for households with one child in rural area. Column (3) shows results for households with no children in urban area. Column (4) shows results for households with no children in rural area. Distributor factors include female income share, age and education difference of husband and wife, and head sex dummy. Demographic information entering utility function include monthly household total income, female age and education, children's age and sex if applicable, labor supply decision, car ownership, house ownership, household points, clothes price, rural fixed effect, monthly fixed effect and state fixed effect.

Table A.5: Effect of female income share on clothing expenditure by north and south region

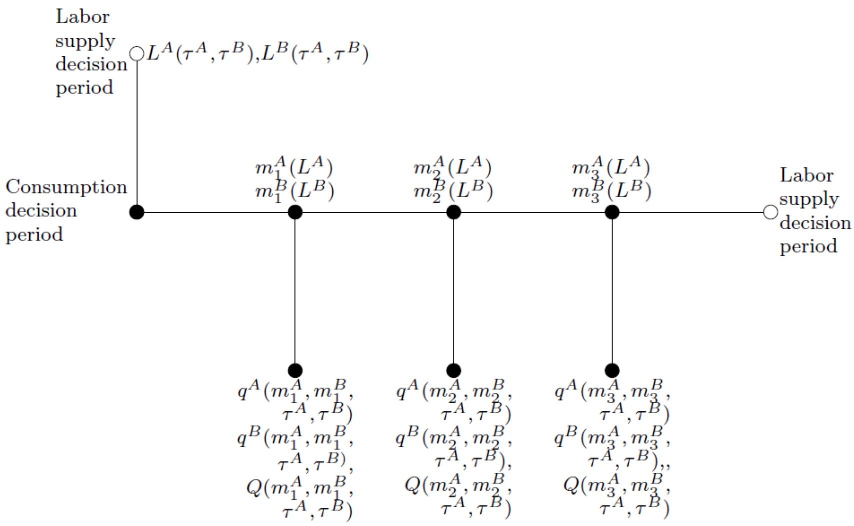
	(1)	(2)	(3)	(4)
	with children		without children	
	South	North	South	North
Panel A: Without control function				
children's clothing budget share	***0.112	0.040		
	0.042	0.047		
female clothing budget share	0.090	0.082	0.094	0.076
	0.059	0.066	0.063	0.058
Panel B: With control function				
children's clothing budget share	0.095	0.025		
	0.101	0.106		
female clothing budget share	0.037	0.034	0.081	0.077
	0.112	0.118	0.115	0.107

Notes: *** p<0.01, ** p<0.05, * p<0.1. The results show effect of female income share on clothing expenditure share. Panel A shows results without controlling individual unobserved characteristics. Panel B shows results control for individual unobserved characteristics. Column (1) shows results for households with one child in South part of Brazil. Column (2) shows results for households with one child in North part of Brazil. Column (3) shows results for households with no children in South part of Brazil. Column (4) shows results for households with no children in North part of Brazil. Distributor factors include female income share, age and education difference of husband and wife, and head sex dummy. Demographic information entering utility function include monthly household total income, female age and education, children's age and sex if applicable, labor supply decision, car ownership, house ownership, household points, clothes price, rural fixed effect, monthly fixed effect and state fixed effect.

Appendix B

Chapter 1: Figures

Figure B.1: Decision Making Process



Notes: This Figure illustrates decision making process for a household. The hollow points denote labor supply decision periods, and the black points denote consumption decision periods. A and B denote individuals. L denotes labor supply. m denotes income. q denotes private goods and Q denotes public goods. τ denotes unobserved characteristics. Firstly, individuals make labor supply decisions which depend on unobserved characteristics. Income of each individual depends on labor supply decisions only. Conditioning on labor supply decisions, income is independent of unobserved characteristics. Secondly. individuals make consumption decisions, which depend on income of each individual and unobserved characteristics. One labor supply decision is followed by several consumption decisions. The Household may reenter labor supply decision period after several consumption decisions.

Appendix C

Chapter 1: Points Calculation for Household Wealth Level

1. Points from bathroom
   - 0 point for no bathroom
   - 1 point for 1 bathroom
   - 2 points for 2 bathrooms
   - 3 points for 3 bathrooms and above
2. Points from households heads education
   - 0 point for not finishing primary school or lower
   - 1 point for finishing primary school but not finishing high school
   - 2 points for finishing high school but not finishing college
   - 3 points for finishing college but not finishing other higher degrees
   - 5 points for finishing degrees higher than college

   If households' head didn't obtain any degree or information for degree is missing, the points will be re-calculated based on years of schooling as below:

- 1 point for years of schooling between 4 years and 8 years
- 2 points for years of schooling between 9 years and 11 years
- 3 points for years of schooling of 12 years and above

3. Points from servant

   Servants that are paid monthly are considered for points calculation.

   - 0 point for no servant
   - 2 points for 1 servant
   - 4 points for 2 servants and above

4. Points from asset holding

   For each asset considered below, additional points are added to households points depending on the number of specific asset.

   Color TVs:

   - 2 points for 1 TV
   - 3 points for 2 TVs
   - 4 points for 3 TVs
   - 5 points for 4 TVs and above

   Radios:

- 1 points for 1 radio
- 2 points for 2 radios
- 3 points for 3 radios
- 4 points for 4 radios and above

Cars:

- 2 points for 1 car
- 4 points for 2 cars
- 5 points for 3 cars and above

Hovers and similar:

- 1 point for 1 hover and above

Laundry machines

- 1 point for 1 Laundry machines and above

Videos including DVD players

- 2 points for 1 video and above

Fridges:

- 2 points for 1 fridge and above

Freezers:

- 2 points for 1 freezer and above

Appendix D

Chapter 1: Derivation

1. **Existence of function ϕ such that $\tau^B = \phi(\tau^A)$.**

   Let $m(\tau^A, \tau^B) = m^A(\tau^A, \tau^B) + m^B(\tau^A, \tau^B)$.

   By Assumption 3, $\frac{\partial m(\tau^A,\tau^B)}{\partial \tau^A} > 0$ and $\frac{\partial m(\tau^A,\tau^B)}{\partial \tau^B} > 0$

   As $m(\tau^A, \tau^B) = \bar{m}$, apply Implicit Function Theorem, there exists a function ϕ such that $\tau^B = \phi(\tau^A)$.

   Moreover, $\frac{\mathbf{d}\tau^B}{\mathbf{d}\tau^A} = -\frac{m(\tau^A,\tau^B)/\partial\tau^A}{m(\tau^A,\tau^B)/\partial\tau^B} < 0$

2. **Derivation of equation (1.5), (1.6) and (1.7).**

   Denote λ as Lagrange multiplier associated with budget constrain. Solution of program (1.4) can be categorized as:

$$\mu(s(\tau^A, \phi(\tau^A)))\frac{\partial u^A(q^A, Q, \tau^A)}{\partial q^A} = \lambda p^A$$
$$[1 - \mu(s(\tau^A, \phi(\tau^A)))]\frac{\partial u^B(q^B, Q, \phi(\tau^A))}{\partial q^B} = \lambda p^B$$

$$\mu(s(\tau^A, \phi(\tau^A)))\frac{\partial u^A(q^A, Q, \tau^A)}{\partial q} +$$
$$[1 - \mu(s(\tau^A, \phi(\tau^A)))]\frac{\partial u^B(q^B, Q, \phi(\tau^A))}{\partial Q} = \lambda p^Q$$

W.O.L.G, assume all the prices are 1. Combined with budget constrain, the above program can be solved. Optimal consumption bundle depends on μ[1], τ^A, $\bar{m}$. Ignore $\bar{m}$ as this can be controlled in empirical work, equation (1.5), (1.6), and (1.7) can be obtained by taking total differentiation.

3. $\frac{\partial q^{A*}}{\partial \mu} > 0$ **for no public good case.**

   Based on Part 2, we can get: $\frac{\partial u^A(q^{A*}, \tau^A)/\partial q^A}{\partial u^B(q^{B*}, \phi(\tau^A))/\partial q^B} = \frac{1-\mu}{\mu}$. As utility function is concave, the above inequality can be obtained.

4. $\frac{\partial q^{A*}}{\partial \tau^A} > 0$ **for no public good case.**

   $\frac{\partial q^{A*}}{\partial \tau^A} = \mu * \frac{\partial u^A(q^{A*}, \tau^A)/\partial \tau^A}{\partial u^A(q^{A*}, \tau^A)/\partial q^A} > 0$

5. $\frac{\mathbf{d}s}{\mathbf{d}\tau^A} > 0$

$$\begin{aligned}\frac{\mathbf{d}s(\tau^A, \phi(\tau^A))}{\mathbf{d}\tau^A} &= \frac{1}{\bar{m}}(\frac{\partial(m^A(\tau^A, \phi(\tau^A)))}{\partial \tau^A} - \frac{\partial(m^A(\tau^A, \phi(\tau^A)))}{\partial \phi}\frac{\partial \phi(\tau^A)}{\tau^A}) \\ &= \frac{1}{\bar{m}}(\frac{\partial(m^A(\tau^A, \phi(\tau^A)))}{\partial \tau^A}) > 0\end{aligned}$$

6. $\frac{\partial Q^*}{\partial \tau^A} > 0$ **for the case only A cares about public good.**

   $\frac{\partial Q^*}{\partial \tau^A} = \mu * \frac{\partial u^A(q^{A*}, Q, \tau^A)/\partial \tau^A}{\partial u^A(q^{A*}, Q, \tau^A)/\partial Q} > 0.$

7. $\frac{\mathbf{d}q^{A*}}{\mathbf{d}\tau^A} > 0$ **for private consumption in second stage of two-stage budgeting.**

[1]Given μ, s does not affect consumption.

Based on Part 3, we can get

$$\frac{\partial^2 u^A(q^{A*},\tau^A)}{\partial q^A \partial q^A}\frac{\mathbf{d}q^{A*}}{\mathbf{d}\tau^A}+\frac{\partial^2 u^A(q^{A*},\tau^A)}{\partial q^A \partial \tau^A}=\frac{1-\mu}{\mu}\frac{\partial^2 u^B(q^{B*},\phi(\tau^A))}{\partial q^B \partial \phi}\frac{\mathbf{d}\phi(\tau^A)}{\mathbf{d}\tau^A}$$

By Assumption 2, the term on right hand side is negative. The second term on left hand side is positive. As $\frac{\partial^2 u^A(q^{A*},\tau^A)}{\partial q^A \partial q^A}$ is negative, $\frac{\mathbf{d}q^{A*}}{\mathbf{d}\tau^A}$ must be positive.

Appendix E

Chapter 2: Tables

Table E.1: Summary statistics

	(1) Married Households	(2) Cohabiting Households	(3) Differences
No. of households	2848	504	
Male parter's age	40.02	37.99	2.02***
Female partner's age	37.15	35.07	2.08***
No. of children	2.73	2.54	0.19***
Log asset value	10.61	9.92	0.69***
Male parter's education High school and above	0.24	0.18	0.06***
Female partner's education High school and above	0.161	0.105	0.056***
Male parter's log income	9.48	9.38	0.10
Female partner's log income	2.84	3.38	-0.54**
Female partner's log income conditioning on working	9.31	9.66	-0.35***
Male parter supplies labor	0.876	0.859	0.017
Female partner supplies labor	0.284	0.315	-0.031
Years of union formation	17.96	13.49	4.47***

Notes: *** p<0.01, ** p<0.05, * p<0.1. This Table shows summary statistics for sample used in chapter 2 for year 2002. Column (1) shows statistics for married households. Column (2) shows statistics for cohabiting households. Column (3) shows difference between two types of households.

Table E.2: Cohabiting households: living arrangement movement

	(1) separated	(2) married
Male education	0.161	0.230
	(0.419)	(0.280)
Female education	-0.981	0.877**
	(0.708)	(0.340)
Log total asset	-0.096*	-0.016
	(0.054)	(0.042)
Number of children	-0.172*	-0.031
	(0.098)	(0.069)
log male yearly income	-0.051	0.025
	(0.049)	(0.040)
log female yearly income	0.085**	-0.002
	(0.033)	(0.024)
Male age	-0.008	0.015
	(0.026)	(0.018)
Female age	-0.013	-0.020
	(0.025)	(0.182)
Male asset share	1.661**	0.360
	(0.660)	(0.457)
Constant	-0.761	-1.160
	(1.310)	(0.929)

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. This Table shows results examining living arrangement movement for cohabiting households in year 2002 based on multinomial logit regression. The baseline outcome is staying *cohabiting*. Column (1) is for the outcome *separated*. Column (2) is for the outcome *married*. The number shows how different demographics affect the probability of moving to difference living arrangement compared with staying cohabiting.

Table E.3: Married households: divorce rate

	(1)
	divorce
Male education	-0.310***
	(0.105)
Female education	0.005
	(0.116)
Log total asset	-0.006
	(0.013)
Number of children	0.010
	(0.025)
Log male yearly income	-0.004
	(0.013)
Log female yearly income	0.020**
	(0.020)
Male age	0.001
	(0.009)
Female age	-0.021**
	(0.009)
Male asset share	0.033
	(0.181)
constant	-0.626
	0.323

Notes: This Table shows results examining divorce behavior of married households in year 2002 using Probit regression. The numbers show the effects of difference demographics on divorce probability.

Table E.4: Preset parameters

Parameter	value	Reference
Initial age	25	
Years in each period	5	
Retirement age	65	
Age at death	75	
RRA (δ_Q)	1.5	Attanasio et al. (2008)
Discount factor (β)	0.98	Attanasio et al. (2008)
Input for household production (α)	0.73	Modified OECD scale
Equivalent scale (e)		Modified OECD scale

Notes: This Tables shows values for preset parameters with corresponding references if applicable.

Table E.5: Moment matching for living arrangement

	(1)	(2)	(3)	(4)
	Low education		High education	
	data	sim data	data	sim data
Panel A: Married Households				
Married	93.01%	93.9%	94.99%	94.93%
Cohabit	3.14%	3.73%	2.23%	2.59%
Divorce	3.85%	2.37%	2.77%	2.47%
Panel B: Cohabiting Households				
Married	26.37%	21.04%	31.63%	23.57%
Cohabit	65.78%	65.78%	62.67%	64.29%
Divorce	7.85%	13.19%	5.61%	12.24%

Notes: This Table shows results for moment of living arrangement corresponds to Figure F.7 and Figure F.8. Panel A is for married households and Panel B is for cohabiting households. Column (1) shows moments for low education households based on Mexican data. Column (2) shows moments for low education households based on simulated data. Column (3) shows moments for high education households based on Mexican data. Column (4) shows moments for high education households based on simulated data.

Table E.6: Results: Parameters for Structural Model

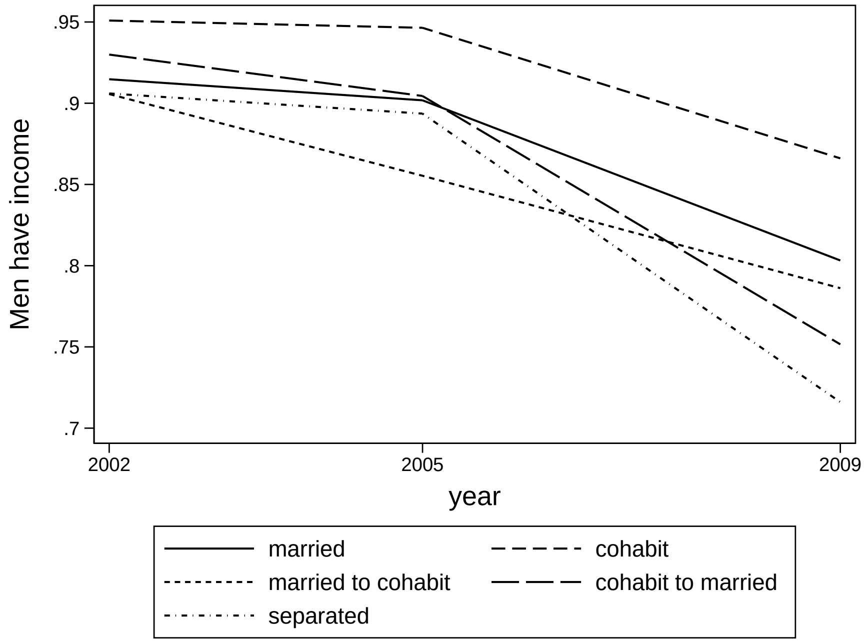
Parameter	Ceof	SE
Panel A: **Low education match households**		
Mean of matching quality for married households (μ_m)	0.17***	(0.013)
SD of matching quality for married households (σ_m)	0.05***	(0.016)
Mean of matching quality for cohabiting households (μ_c)	0.15***	(0.021)
SD of matching quality for cohabiting households (σ_c)	0.25***	(0.040)
Panel B: **High education match households**		
Mean of matching quality for married households (μ_m)	0.26*	(0.137)
SD of matching quality for married households (σ_m)	0.05	(0.038)
Mean of matching quality for cohabiting households (μ_c)	0.21***	(0.007)
SD of matching quality for cohabiting households (σ_c)	0.30***	(0.115)
Panel C: **Common parameters**		
Utility from leisure (δ_l)	$1.2*10^{-6***}$	$(2.6*10^{-8})$
Penalty for not working for men (ϕ)	0.43***	(0.084)

Notes: This Table shows results for parameters in structural model. Panel A is for low education households. Panel B is for high education households. Panel C shows parameters common across high education and low education households.

Appendix F

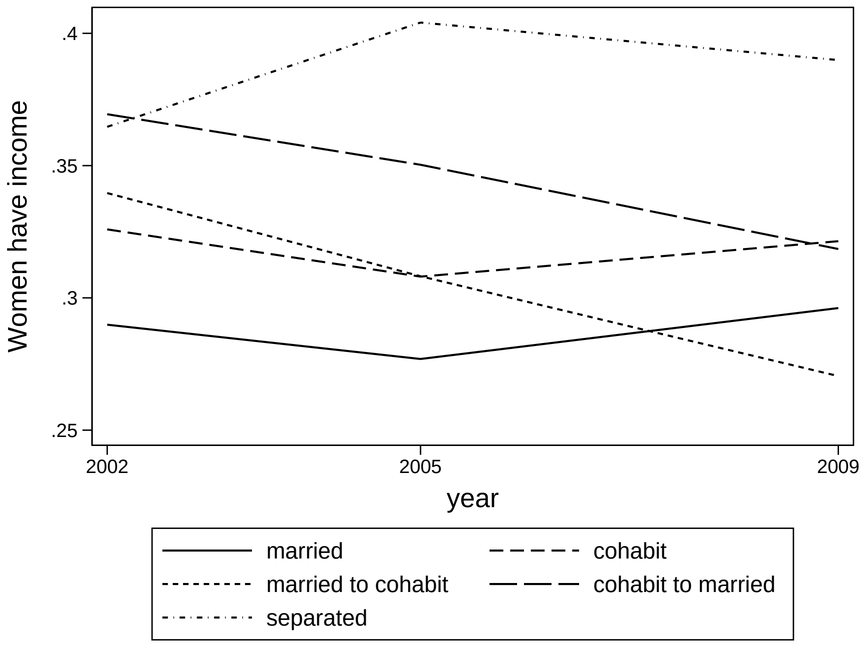
Chapter 2: Figures

Figure F.1: Men: have income



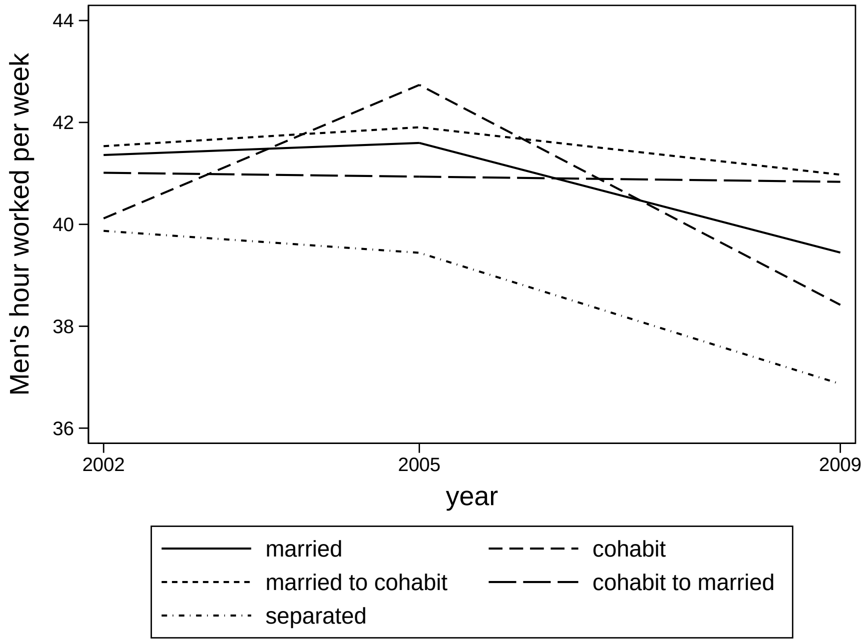
Notes: This Figure shows proportion of men that have income based on five types of households. The solid line represents households being married for three panels. The dash line represents households being cohabiting for three panels. The short dash line represents households that moved from marriage to cohabitation. The long dash line represents households that moved from cohabitation to marriage. The dash-dot line represents households that got divorced or separated.

Figure F.2: Women: have income



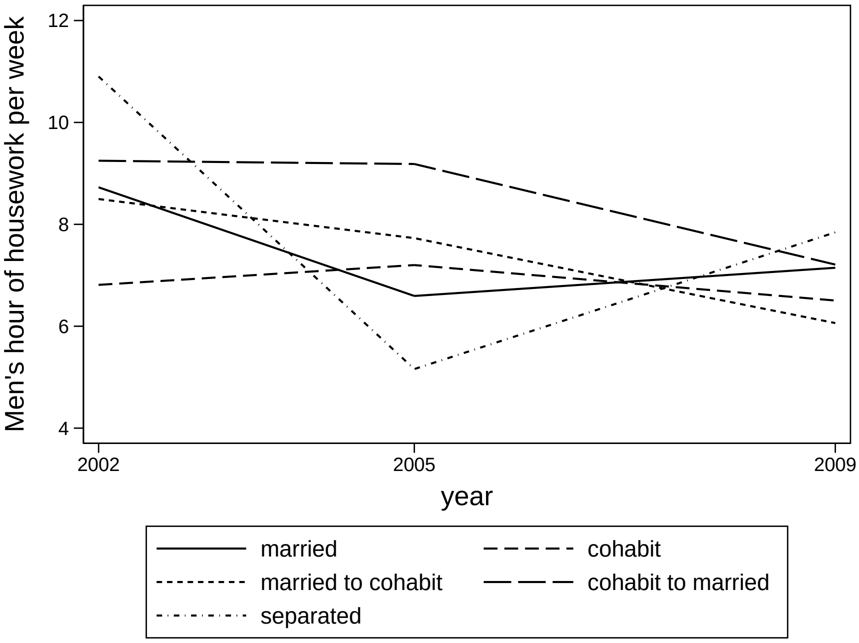
Notes: This Figure shows proportion of women that have income based on five types of households. The solid line represents households being married for three panels. The dash line represents households being cohabiting for three panels. The short dash line represents households that moved from marriage to cohabitation. The long dash line represents households that moved from cohabitation to marriage. The dash-dot line represents households that got divorced or separated.

Figure F.3: Men: hours worked per week



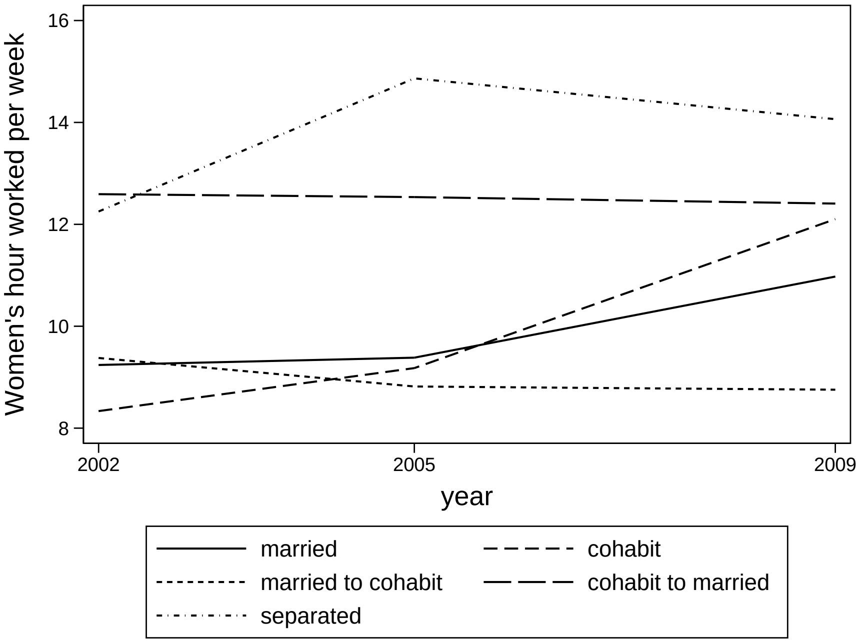
Notes: This Figure shows average hour of work per week for men based on five types of households. The solid line represents households being married for three panels. The dash line represents households being cohabiting for three panels. The short dash line represents households that moved from marriage to cohabitation. The long dash line represents households that moved from cohabitation to marriage. The dash-dot line represents households that got divorced or separated.

Figure F.4: Men: hours of housework per week



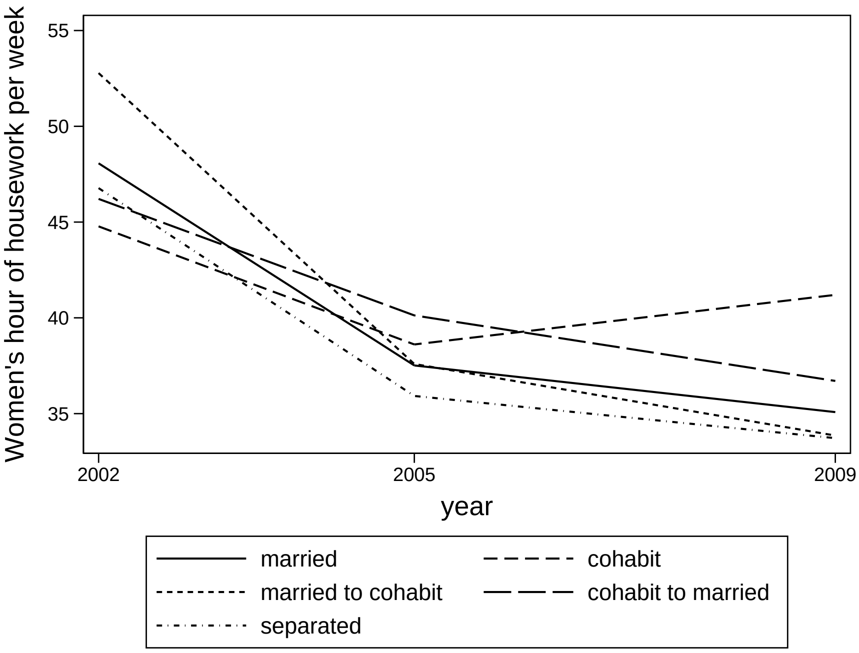
Notes: This Figure shows average hour of housework per week for men based on five types of households. The solid line represents households being married for three panels. The dash line represents households being cohabiting for three panels. The short dash line represents households that moved from marriage to cohabitation. The long dash line represents households that moved from cohabitation to marriage. The dash-dot line represents households that got divorced or separated.

Figure F.5: Women: hours worked per week



Notes: This Figure shows average hour of work per week for women based on five types of households. The solid line represents households being married for three panels. The dash line represents households being cohabiting for three panels. The short dash line represents households that moved from marriage to cohabitation. The long dash line represents households that moved from cohabitation to marriage. The dash-dot line represents households that got divorced or separated.

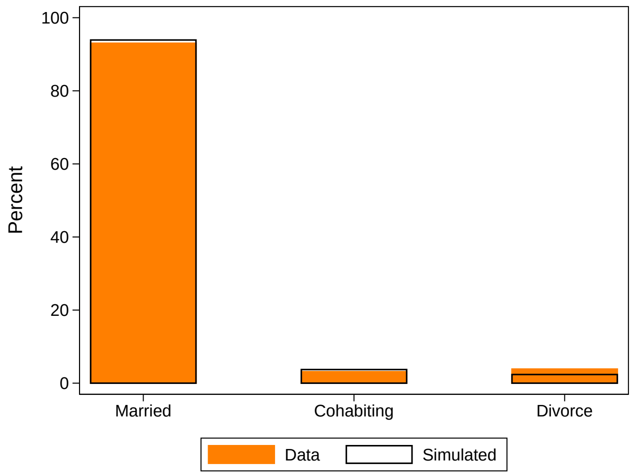
Figure F.6: Women: hours of housework per week



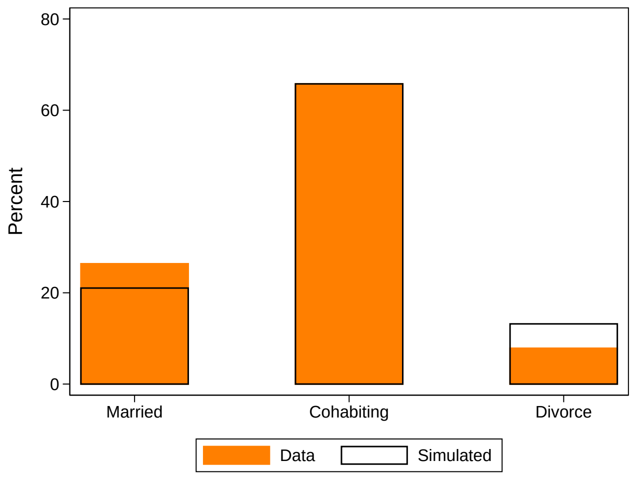
Notes: This Figure shows average hour of housework per week for women based on five types of households. The solid line represents households being married for three panels. The dash line represents households being cohabiting for three panels. The short dash line represents households that moved from marriage to cohabitation. The long dash line represents households that moved from cohabitation to marriage. The dash-dot line represents households that got divorced or separated.

Figure F.7: Moment of matching: living arrangement

(a) Married: low education



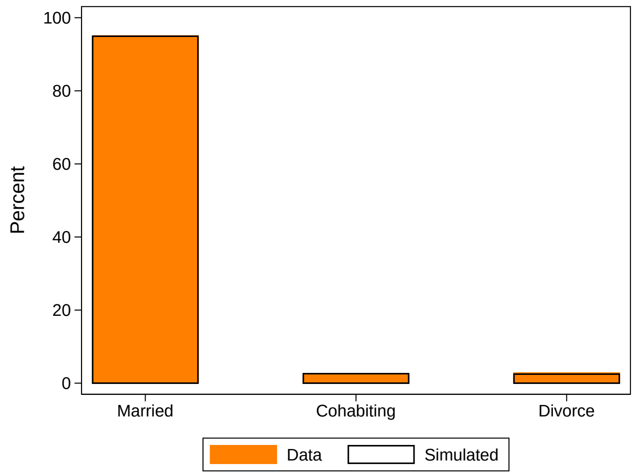
(b) Cohabiting: low education



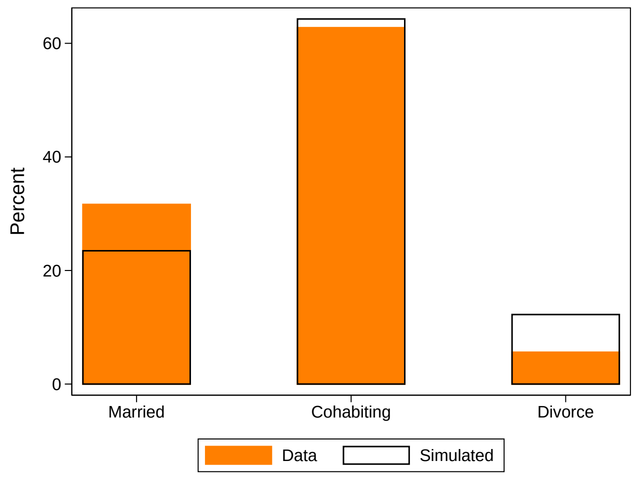
Notes: This Figure shows matching of moments for living arrangements of low education households. Panel a is for married households and Panel b is for cohabiting households. The colored bars are based on Mexican data. The empty bars are based on simulated data.

Figure F.8: Moment of matching: living arrangement

(a) Married: high education



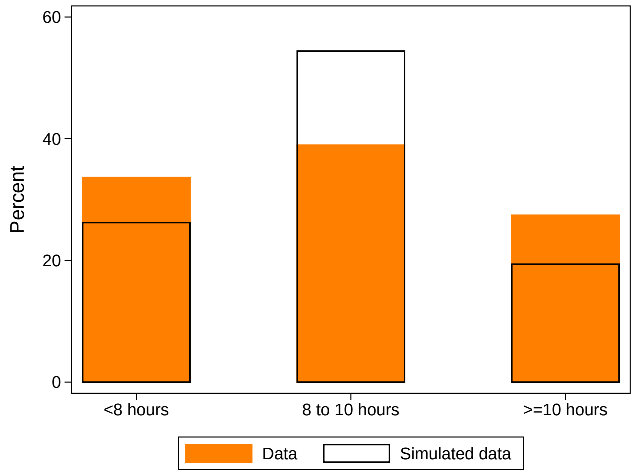
(b) Cohabiting: high education



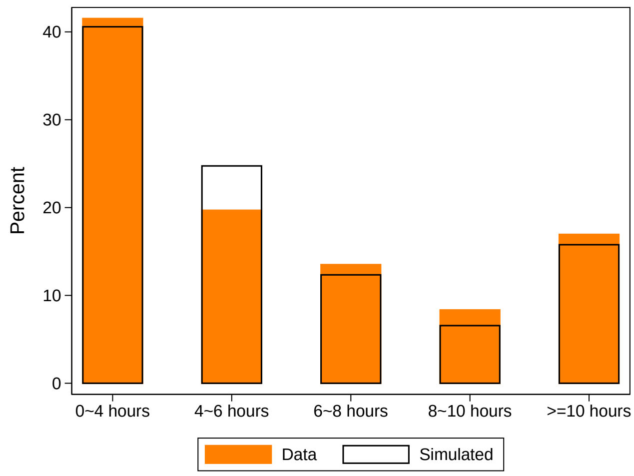
Notes: This Figure shows matching of moments for living arrangements of high education households. Panel a is for married households and Panel b is for cohabiting households The colored bars are based on Mexican data. The empty bars are based on simulated data.

Figure F.9: Moment of matching: male hour of work and female hour of housework

(a) Male hour of work



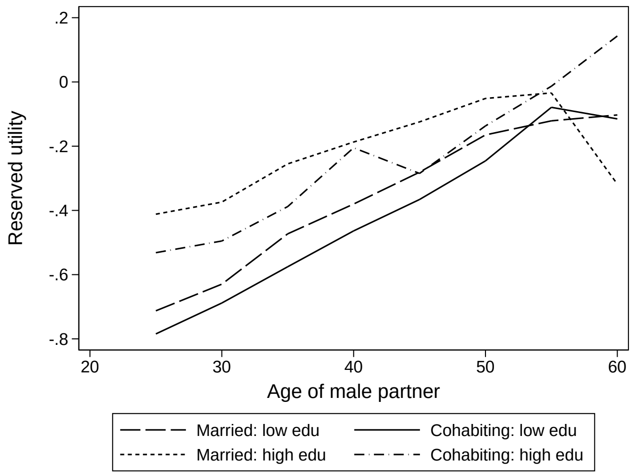
(b) Female hour of housework



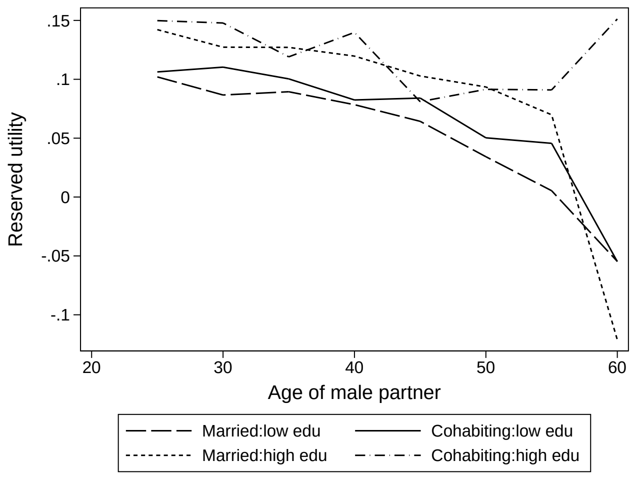
Notes: This Figure shows matching of moments for male labor supply and female housework. Panel a shows histogram of male hour of work per day. Panel b shows histogram of female hour of housework per day. The colored bars are based on Mexican data. The empty bars are based on simulated data.

Figure F.10: Outside options

(a) Female outside options



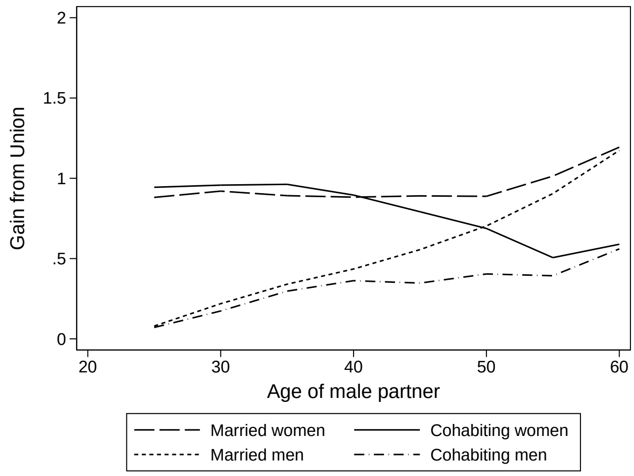
(b) Male outside options



Notes: This Figure shows individual outside options, the utilities individuals can get upon divorce or separation. Panel a is for women and Panel b is for men. The long dash line represents individuals from low education married households. The solid line represents individuals from low education cohabiting households. The short dash line represents individuals from high education married households. The dash line represents individuals from high education cohabiting households.

Figure F.11: Gain from union

(a) Low education

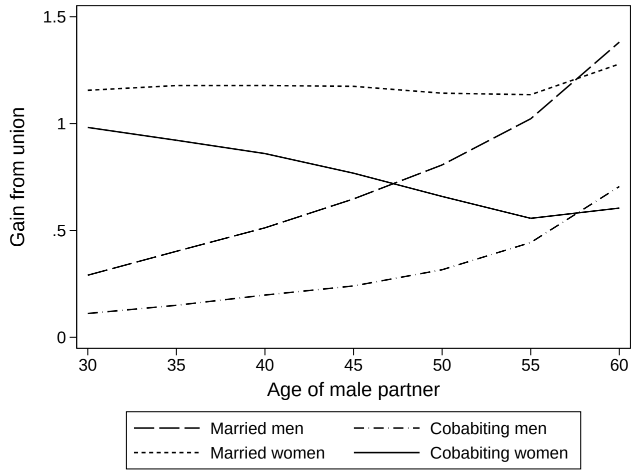


(b) High education

Notes: This Figure shows individuals' gain from forming union based on Mexican data. Panel a is for low education households and Panel b is for high education households. The long dash line represents married women. The solid line represents cohabiting women. The short dash line represents married men. The dash line represents cohabiting men.

Figure F.12: Gain from union

(a) Low education

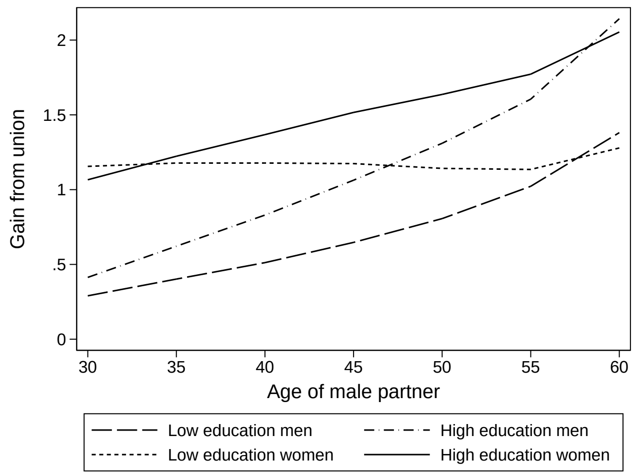


(b) High education

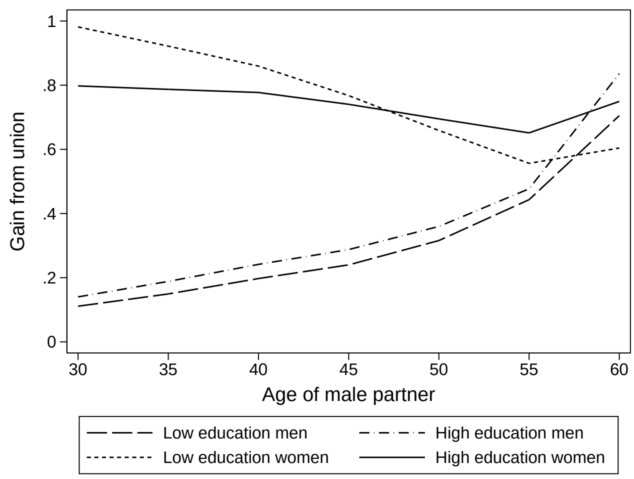
Notes: This Figure shows individuals' gain from forming union based on simulation of representative households. Two households are considered. One for high education and one for low education. The households are assumed to be either married or cohabiting initially. Panel a is for low education representative household and Panel b is for high education representative household. The long dash line represents married men. The dash line represents cohabiting men. The short dash line represents married women. The solid line represents cohabiting women.

Figure F.13: Gain from union

(a) Married households



(b) Cohabiting households



Notes: This Figure shows individuals' gain from forming union based on simulation of representative households. Two households are considered. One for high education match and one for low education match. The households are assumed to be either married or cohabiting initially. Panel a is for married household and Panel b is for cohabiting household. The long dash line represents low education men. The dash line represents high education men. The short dash line represents low education women. The solid line represents high education women.

Appendix G

Chapter 3: Robustness

Appendix Tables H.8 to H.13 (robustness tests) and Appendix Figures I.8 to I.11 (variation in meteorology and road congestion around each discontinuity) are referenced directly in the text.

Here we complement the graphical analysis of Figure I.2 by formally testing whether consumer prices for E20/E25 gasoline fuel, as well as E100 ethanol fuel, were discontinuous at cutoff dates for the first three discontinuities in the composition of gasoline—we take these cutoff points as January 31, 2010 (last day of E25 shipments), May 1, 2010 (first day of E25 shipments), and September 30, 2011 (last day of E25 shipments). (Data are missing from June 2013, thus we exclude discontinuity # 4, on May 1, 2013, from the analysis.) As in our preferred specification for the analysis of pollution, we adopt local linear regression, and examine each discontinuity separately. Overall, we fail to reject that fuel prices varied smoothly at these cutoffs, as suggested by

Figure I.2, but regression estimates are less robust compared to the analysis of ozone concentrations, possibly because we do not observe fuel prices every day.

Our data consists of repeated cross-sections of about 350 fuel retail stations in Sao Paulo city, available from the fuel market regulator (ANP). Enumerators surveyed stations from Monday to Thursday each week, and any given station is visited at most once in a weekly survey. We observe station-level E20/E25 gasoline and E100 ethanol prices, in nominal BR$ per liter, for the regular grade that is dominant for both gasoline and ethanol (midgrade gasoline is typically available at a 4% markup over regular gasoline, but the penetration is low; Salvo and Huse 2013). Despite no data being available for Fridays and weekends, we are fortunate that the data density is fairly balanced around each cutoff point.

In a first-step demeaning procedure, we again consider a sample of 90 days on either side of a cutoff date and regress station-date level fuel prices (E20/E25 gasoline or E100 ethanol, separately) on an eighth-order polynomial trend, 38 station brand fixed effects (e.g., Shell, Esso), and location fixed effects, namely 272 "3-digit" zipcodes (e.g., 056 as in the complete zipcode 05610-060).

Local linear regression estimates in a second step are shown in Table H.12.

As with the analysis of pollution, we normalize the reporting of price effects to correspond to an increase in the ethanol fraction in gasoline fuel, from E20 to E25. Panel A reports on gasoline prices, whereas panel B reports on ethanol prices. Again, we consider both the CCT and IK bandwidth criteria, with CCT invariably selecting shorter bandwidths compared to IK. Selected bandwidths tend to shorter for the analysis of ethanol prices compared to that for gasoline prices. We implement each local linear regression—separately by fuel (across panels), by discontinuity (across columns), by bandwidth criterion (within panels)—at two levels of aggregation: (i) directly on the disaggregate individual station residual prices from the first step; and (ii) taking the mean of these residual prices by 2-digit zipcode and station brand within date, to "cluster" our inference at the zipcode-brand level (there are 476 "2-digit" zipcode by station brand combinations).

We find that both levels of aggregation yield similarly small and statistically insignificant estimates for the effect of changes in the gasoline blend on gasoline and ethanol prices at the pump. Across 12 regression samples/variants for each fuel, the average price effect is -0.01 BR$/liter (about half a US dollar cent) for E20/E25 gasoline, and 0.00 BR$/liter (though somewhat more variable) for E100 ethanol.

Salvo-Geiger's design applied to a longer sam-

ple

Salvo and Geiger (2014) fit empirical models of the form:

$$y_{iht} = fuelmix_t'\lambda + W_{iht}'\Delta^W + T_{iht}'\Delta^T + \nu_{ih} + \mu_{it} + \epsilon_{iht}$$

in which, in a second step, ozone mass concentration y_{iht} at site i, hour h, and date t is regressed on the fuel mix, in particular, the ethanol share in the bi-fuel fleet, imputed from a discrete-choice demand model estimated in a first step (e.g., a multinomial probit, Figure I.1(b)). Meteorological conditions W_{iht}, local road traffic conditions T_{iht}, and time-by-location fixed effects (ν_{ih}, μ_{it}) control for temporal and spatial determinants of ozone pollution other than the fuel mix. The identifying assumption is that, conditional on controls, the residual is uncorrelated with the fuel mix, in particular, $E\left[fuelmix_t\epsilon_{iht}|W_{iht}, T_{iht}, \nu_{ih}, \mu_{it}\right] = 0$.

Table H.13, columns 1 and 4 report estimates from the second-step of the Salvo-Geiger model using an expanded 55-month sample period, November 2008 to May 2013 (rather than to May 2011), and the same 12 monitoring sites across the metropolis. An observation is a date-site pair. (See Section 3 for a description of the data.) The dependent variable is the mean ozone concentration recorded in the afternoon hours between 12pm to 4pm, in μg/m^3 (column 1) or its logarithm (column 4). In column 1, the ozone variation that can be attributed to the in-sample range for the share of bi-fuel vehicle owners choosing E100 ethanol, between 24% and 89%, is estimated at $21.6 \times (.89 - .24) = 14.0$ μg/m^3, with estimated standard error (s.e.) of 4.8 μg/m^3. For perspective, 14.0 μg/m^3 is equivalent to about 7 ppb, or 10% of the 8-hour standards set both for Sao Paulo state (140 μg/m^3) and the US (70 ppb). A 14.0 μg/m^3 rise in ozone from consumers switching into ethanol at the pump amounts to 20% of the sample mean ozone reading of 69.4 μg/m^3. Standard errors are obtained through a bootstrap procedure to account for sampling error in the first-step estimation of the ethanol choice probability in the bi-fuel fleet—see Table H.13 notes. Ozone levels rise in contemporaneously recorded radiation and temperature, and fall in humidity and wind speed. Omitted from the table for brevity, ozone levels also tend to rise when thermal inversions are recorded hours earlier, at 9am, and with increased traffic congestion in the morning commuting hours (7am to 11am). (See Lin et al. 2001 for references on the meteorological dependence of ozone formation, including biogenic hydrocarbon emissions.) Estimates are robust to excluding the colder months of June to September from the sample, and restricting the sample to non-holiday weekdays, among several other variations.

As a first variant to this bootstrap procedure, we employ two-stage least squares (2SLS) with the observed ethanol-to-gasoline price ratio (Figure I.1(a)) instrumenting for the predicted ethanol share in the bi-fuel fleet. Salvo and Huse (2013) and Salvo (2016) observed consumers making choices at the pump between 2010 and 2012. They find (see Figs. 4-6 and Figs. 3-4, respectively) that the relationship between the estimated gasoline choice probability (equivalently, the ethanol choice probability)—an object subject to sampling error—and the ethanol-to-gasoline price ratio in the data—the instrumental variable (IV)—is quite linear over a wide range of price variation. Table H.13 columns 2 and 5 (ozone in levels and logs) show that the 2SLS estimator yields similar estimates to columns 1 and 4, if somewhat higher coefficients and lower standard errors.

A second alternative to using an estimated fuel share in the second-step ozone regression is to use an *observed* fuel share, were this available. Unfortunately, high-frequency (daily) fuel quantity data are not available for the Sao Paulo metropolis. However, we obtained ethanol and blended gasoline shipments reported on a monthly basis by distributors for the state of Sao Paulo over the sample period (again, see the data section). In spite of its lower frequency—monthly variation versus daily—and aggregate geographic coverage—state-level versus metropolis—we use such data in columns 3 and 6 to proxy for the ethanol share (again, with ozone in levels and logs). A similar result obtains, noting that the estimated coefficient is larger in magnitude relative to the other columns as the ethanol share now pertains to the entire population of light vehicles and motorcycles, not only to bi-fuel engines—this aggregate distributor ethanol share is shown in Figure I.1(c).

Appendix H

Chapter 3: Tables

Table H.1: Discontinuities in the composition of retail gasoline fuel, 2010 to 2013, per central government mandate

	Gasoline blend discontinuity			
2-5	# 1	# 2	# 3	# 4
Shift in gasoline composition:	E25 to E20	E20 to E25	E25 to E20	E20 to E25
For distributor shipments starting on:	February 1, 2010	May 1, 2010	October 1, 2011	May 1, 2013
Date cutoff for sharp design:	February 4, 2010	May 4, 2010	October 4, 2011	May 4, 2013
For robustness, assumed phase-in over three or five days from the sharp-design date cutoff:	February 4 to 6, or February 4 to 8, 2010	May 4 to 6, or May 4 to 8, 2010	October 4 to 6, or October 4 to 8, 2011	May 4 to 6, or May 4 to 8, 2013
Season of the discontinuity:	Summer	Fall	Spring	Fall
Aggregate state distributor shipments:				
Reported for the month of:	February 2010	May 2010	October 2011	May 2013
Gasoline shipments (E20 or E25, 10^6 m^3)	654.3	582.0	802.2	927.1
Ethanol shipments (E100, 10^6 m^3)	465.3	751.7	533.7	431.3
Gasoline (E20 or E25) share of VMT	68%	53%	69%	76%
Ethanol (E100) share of VMT	32%	47%	31%	24%
E100 to E20/E25 price ratio at the pump:	0.73	0.55	0.70	0.69

Notes: Changes to the composition of retail gasoline fuel per Ministry of Agriculture Ordinance 7, dated January 11, 2010 (discontinuities # 1 and # 2); Ordinance 678-2011, dated August 31, 2011 (discontinuity # 3); and Ordinance 105-2013, dated February 28, 2013 (discontinuity # 4). We convert fuel volumes (E20, E25, E100) to vehicle miles traveled (VMT) based on mean fuel economy on an urban driving cycle for a sample of 720 bi-fuel vehicles sampled in 2010 (Salvo and Huse 2013), namely, 10.42 km/liter for E20, 10.23 km/liter for E25 and 7.05 km/liter for E100. Fuel volume and price statistics are based on ANP data: volumes at state level by month, prices at pump level by date (shown are means over dates within month of medians across retail stations sampled by date).

Table H.2: Descriptive statistics for the four 180-day samples pooled together

Descriptive statistics for the four 180-day samples pooled together

Variables	(1) N	(2) Mean	(3) Std.Dev.	(4) Min.	(5) Max.
O_3 concentration ($\mu g/m^3$, hourly readings 7am to 11am)	43,903	22.68	20.96	0	179
O_3 concentration ($\mu g/m^3$, hourly readings 12pm to 4pm)	44,228	66.29	36.68	0	353
E25, not E20, gasoline (yes=1)	44,228	0.50	0.50	0	1
PM2.5 concentration ($\mu g/m^3$, hourly readings 7am to 11am)	3,307	19.96	13.50	0	97
PM2.5 concentration ($\mu g/m^3$, hourly readings 5pm to 8pm)	2,607	18.60	11.27	0	86
CO concentration (ppm, hourly readings 7am to 11am)	41,833	1.06	0.71	0	7
CO concentration (ppm, hourly readings 5pm to 8pm)	33,921	1.01	0.71	0	10.22
NOx concentration (ppb, hourly readings 7am to 11am)	39,226	74.47	70.84	0	701
NOx concentration (ppb, hourly readings 5pm to 8pm)	31,843	52.06	54.68	0	588
Temperature (^{0}C, hourly readings 7am to 11am)	3,599	19.72	4.19	6.15	32.37
Temperature (^{0}C, hourly readings 12pm to 4pm)	3,600	24.97	4.77	7.69	35.97
Radiation (W/m^2, hourly readings 7am to 11am)	3,599	240.66	219.72	0	893
Radiation (W/m^2, hourly readings 12pm to 4pm)	3,600	513.66	241.22	13.20	1128.50
Relative humidity (%, hourly readings 7am to 11am)	3,599	80.35	13.92	26	97.98
Relative humidity (%, hourly readings 12pm to 4pm)	3,600	58.56	17.96	13.50	98
Wind speed (m/s, hourly readings 7am to 11am)	3,599	1.37	0.67	0.02	3.99
Wind speed (m/s, hourly readings 12pm to 4pm)	3,600	2.04	0.67	0.28	5.20
Precipitation (mm/h, hourly readings 7am to 11am)	3,592	0.12	0.79	0	16.80
Precipitation (mm/h, hourly readings 12pm to 4pm)	3,597	0.22	1.65	0	43.80
Wind blows from N-E (yes=1, hourly readings 7am to 11am)	13,385	0.28	0.45	0	1
Wind blows from N-W (yes=1, hourly readings 7am to 11am)	13,385	0.22	0.41	0	1
Wind blows from S-W (yes=1, hourly readings 7am to 11am)	13,385	0.07	0.26	0	1
Wind blows from S-E (yes=1, hourly readings 7am to 11am)	13,385	0.33	0.47	0	1
Wind blows from N-E (yes=1, hourly readings 12pm to 4pm)	13,379	0.14	0.34	0	1
Wind blows from N-W (yes=1, hourly readings 12pm to 4pm)	13,379	0.41	0.49	0	1
Wind blows from S-W (yes=1, hourly readings 12pm to 4pm)	13,379	0.11	0.31	0	1
Wind blows from S-E (yes=1, hourly readings 12pm to 4pm)	13,379	0.34	0.47	0	1
Thermal inversion at 9am, base 0-200m (yes=1)	718	0.09	0.29	0	1
Thermal inversion at 9am, base 200-500m (yes=1)	718	0.27	0.44	0	1
Thermal inversion at 9pm, base 0-200m (yes=1)	717	0.25	0.43	0	1
Thermal inversion at 9pm, base 200-500m (yes=1)	717	0.12	0.33	0	1
Road congestion (km, citywide, hourly readings 7am to 11am)	3,600	35.32	33.62	0	156.38
Road congestion (km, citywide, hourly readings 5pm to 8pm)	2,880	57.33	50.05	0	277.20
Per-liter ethanol-to-gasoline price ratio	720	66.35	5.86	52.63	74.13

Notes: The sample consists of 90 days on either side of the cutoff date for each of the four discontinuities described in Table H.1. For pollutant concentrations, an observation is a date-hour by site (within the indicated hour window). For the E25 gasoline indicator, an observation is a date-hour by ozone site (within the 12pm to 4pm window) for which ozone concentrations were recorded, with the purpose of demonstrating that ozone measurements are similarly available on either side of the cutoff dates. For meteorology and citywide road congestion, an observation is a date-hour (we average across weather stations), except wind direction, where an observation is a date-hour by site (there are up to five sites measuring wind direction). Wind direction is interacted with a dummy variable indicating that wind is blowing at a speed in excess of 0.5 m/s. For the fuel price ratio, an observation is a date. See the text for sources.

Table H.3: OLS regression for Ozone using the 180-day sample, with an eighth-order polynomial trend over the whole period (afternoon hours)

	Gasoline blend discontinuity			
2-5	# 1	# 2	# 3	# 4
	Summer 2010	Fall 2010	Spring 2011	Fall 2013
Dependent variable: Log ozone concentration, by site-hour-date				
E25, not E20, gasoline (yes=1)	0.110**	0.157***	0.117***	0.116***
i.e., higher ethanol content	(0.043)	(0.050)	(0.036)	(0.039)
Control variables (all effects by site-hour)				
Site by hour fixed effects interacted with:				
Eighth-order polynomial time trend	Yes	Yes	Yes	Yes
Day-of-week fixed effects	Yes	Yes	Yes	Yes
Meteorological conditions	Yes	Yes	Yes	Yes
Thermal inversion indicators	Yes	Yes	Yes	Yes
Road traffic congestion	Yes	Yes	Yes	Yes
Ethanol-to-gasoline price ratio (cubic)	Yes	Yes	Yes	Yes
R^2	82%	86%	84%	83%
Number of observations	9,852	10,056	10,122	13,487
Number of regressors	2,161	2,161	2,101	2,976
Number of sites	12	12	12	17

Notes: *** p<0.01, ** p<0.05, * p<0.1. An observation is a site-hour-date triple. The sample includes all hours between 12pm and 4pm on all dates that are plus and minus 90 days of the cutoff point (February4, 2010 for # 1; May 4, 2010 for # 2; October 4, 2011 for # 3; May 4, 2013 for # 4) for all monitoring sites that are active in the Sao Paulo metropolitan area (by active we require at least 70% of a maximum of $180 \times 5 = 900$ hourly measurements to be available in a 180-day period). The dependent variable is the logarithm of hourly ozone concentration ($\mu g/m^3$). Meteorological conditions include quadratics in contemporaneous log temperature, log radiation, log humidity and log wind speed; indicators for contemporaneous wind blowing in from each of four wind direction quadrants; mean precipitation in the contemporaneous hour to three hours earlier. Other covariates include indicators for the base of an atmospheric thermal inversion layer lying within 200m, or between 200 and 500m, from the ground, according to the 9am or 9pm reading, whichever is closest to the given hour; the log of citywide extension of traffic congestion from 7am to 11am interacted with an indicator for the date being a weekday and non-public holiday; a cubic in the ethanol-to-gasoline price ratio lagged by 4 days. All controls enter the specification interacted with site by hour fixed effects. OLS estimates for four separate regressions. Standard errors, in parentheses, are clustered by site-date pair.

Table H.4: Local linear regression for Mean Ozone and Maximum Ozone (afternoon hours)

	Gasoline blend discontinuity			
2-5	# 1	# 2	# 3	# 4
	Summer 2010	Fall 2010	Spring 2011	Fall 2013
A: Dependent variable: Mean (of demeaned) log ozone concentration, by site-date				
CCT bandwidth criterion				
E25, not E20, gasoline (yes=1)	0.070**	0.112**	0.126***	0.078***
i.e., higher ethanol content	(0.032)	(0.045)	(0.036)	(0.029)
Bandwidth	30	25	22	28
Number of observations	642	550	482	883
IK bandwidth criterion				
E25, not E20, gasoline (yes=1)	0.066**	0.105***	0.118***	0.041
i.e., higher ethanol content	(0.032)	(0.037)	(0.035)	(0.032)
Bandwidth	31	35	24	23
Number of observations	663	778	526	726
Number of sites	12	12	12	17
B: Dependent variable: Demeaned maximum of log ozone concentration, by site-date				
CCT bandwidth criterion				
E25, not E20, gasoline (yes=1)	0.102***	0.094**	0.054	0.080**
i.e., higher ethanol content	(0.033)	(0.046)	(0.039)	(0.036)
Bandwidth	27	27	26	29
Number of observations	568	594	619	1,001
IK bandwidth criterion				
E25, not E20, gasoline (yes=1)	0.091***	0.096**	0.054	0.079**
i.e., higher ethanol content	(0.032)	(0.043)	(0.039)	(0.036)
Bandwidth	32	31	26	28
Number of observations	679	685	619	965
Number of sites	12	12	12	17

Notes: *** p<0.01, ** p<0.05, * p<0.1. An observation in the second step is a site-date pair within the bandwidth (in days from the cutoff point) specified by either the CCT or IK selection criterion, for all monitoring sites that are active in the Sao Paulo metropolitan area. See the text for the first-step procedure: to demean hourly log pollutant concentrations (12pm to 4pm) using each 180-day sample, and to collapse residual concentrations to the site-date level taking the mean across hours (panel A); or to take the maximum log pollutant concentration between 12pm and 4pm, and to demean these site-date varying maximum values using each 180-day sample (panel B). See the notes to Table H.3 for site-hour (site in panel B) specific eighth-order polynomial trend, day-of-week fixed effects, meteorology, thermal inversion, road traffic congestion and fuel price controls used to demean the data. Local linear regression estimates for 16 separate regressions. Robust standard errors in parentheses.

Table H.5: Local linear regression for Mean Ozone (afternoon hours): Sensitivity to pooling across the four 180-day samples

2-5	Gasoline blend discontinuity # 1, # 2, # 3, # 4 pooled together, stacked by day plus or minus from cutoff
Dependent variable: Mean (of demeaned) log ozone concentration, by site-date	
CCT bandwidth criterion	
E25, not E20, gasoline (yes=1)	0.081***
i.e., higher ethanol content	(0.017)
Bandwidth	23
Number of observations	2,221
IK bandwidth criterion	
E25, not E20, gasoline (yes=1)	0.083***
i.e., higher ethanol content	(0.016)
Bandwidth	27
Number of observations	2,619
Number of site-discontinuity pairs	12 + 12 + 12 + 17 = 53

Notes: *** p<0.01, ** p<0.05, * p<0.1. An observation in the second step is a site-date pair within the bandwidth (in days from the cutoff point) specified by either the CCT or IK selection criterion, for all monitoring sites that are active in the Sao Paulo metropolitan area. The first step (demeaning hourly log ozone concentrations separately by 180-day sample and taking means across hours within site-date) follows Table H.4A exactly. Local linear regression estimates for two separate regressions. Robust standard errors in parentheses.

Table H.6: Local linear regression for Mean Ozone: Sensitivity to hour-of-the-day and functional form

2-5	Gasoline blend discontinuity # 1 Sum. '10	# 2 Fall '10	# 3 Spr. '11	# 4 Fall '13	All four **pooled** together
A: Dependent variable: Mean (of demeaned) log ozone concentration, 7am to 11am					
CCT **E25, not E20, gasoline** (yes=1)	0.121	0.187**	0.165***	0.084	0.121***
i.e., higher ethanol content	(0.118)	(0.089)	(0.057)	(0.054)	(0.040)
Bandwidth	27	24	34	36	25
IK **E25, not E20, gasoline** (yes=1)	0.093	0.182**	0.175***	0.043	0.121***
i.e., higher ethanol content	(0.107)	(0.087)	(0.059)	(0.062)	(0.038)
Bandwidth	33	25	32	29	27
B: Dep. var.: Mean (of demean.) log ozone conc., 11am to 6pm (modal 8-hour period for max.)					
CCT **E25, not E20, gasoline** (yes=1)	0.054	0.160***	0.142***	0.033	0.072***
i.e., higher ethanol content	(0.035)	(0.046)	(0.039)	(0.030)	(0.018)
Bandwidth	30	24	21	28	24
IK **E25, not E20, gasoline** (yes=1)	0.051	0.149***	0.094***	0.026	0.074***
i.e., higher ethanol content	(0.034)	(0.040)	(0.033)	(0.031)	(0.017)
Bandwidth	33	32	32	27	27
C: Dep. var.: Mean (of demeaned) log ozone concentration, maximum 8-hour average (realized)					
CCT **E25, not E20, gasoline** (yes=1)	0.080*	0.156***	0.060*	0.066**	0.072***
i.e., higher ethanol content	(0.042)	(0.047)	(0.035)	(0.032)	(0.019)
Bandwidth	31	24	25	32	26
IK **E25, not E20, gasoline** (yes=1)	0.072*	0.095***	0.068**	0.029	0.072***
i.e., higher ethanol content	(0.041)	(0.031)	(0.032)	(0.037)	(0.019)
Bandwidth	34	56	32	25	26
D: Dep. var.: Mean (of demeaned) ozone concentration, maximum 8-hour average (realized)					
CCT **E25, not E20, gasoline** (yes=1)	10.99***	2.06	6.52**	3.74*	4.74***
i.e., higher ethanol content	(3.04)	(2.18)	(2.69)	(2.03)	(1.23)
Bandwidth	23	35	34	26	27
IK **E25, not E20, gasoline** (yes=1)	11.07***	1.80	6.37**	3.47*	4.79***
i.e., higher ethanol content	(2.67)	(1.97)	(3.21)	(1.90)	(1.20)
Bandwidth	28	41	24	29	28
Mean value dep.var. (max 8-h avg., $\mu g/m^3$)	62.07	55.70	72.28	52.24	59.85
No. of sites (site-discontinuity pairs)	12	12	12	17	53

Notes: *** p<0.01, ** p<0.05, * p<0.1. All notes to Tables H.4 and H.5 (respectively, second-step regressions are separate by discontinuity or pooled across discontinuities) apply here. Instead of afternoon hours (12pm to 4pm), panels A and B consider morning hours (7am to 11am) and the modal (most common) 8-hour period for daily maximum 8-hour averages (11am to 6pm), respectively. Panel C considers the realized 8-hour period, by site-date in each 180-sample, of daily maximum 8-hour averages. Since this realized 8-hour period varies from day to day (e.g., 10am to 5pm versus 11am to 6pm), we interact first-step controls only with site fixed effects, as in Table H.10, with the exception of day-of-week which we interact with site-hour fixed effects. Panel D follows the third panel but considers (the mean of demeaned) ozone concentrations rather than log ozone concentrations as in Tables H.4 and H.5. Local linear regression estimates for 40 separate regressions. For brevity, we omit the number of observations (these scale with the tested bandwidth). Robust standard errors in parentheses.

Table H.7: Local linear regression for Mean PM2.5, Mean CO and Mean NOx (evening hours)

2-5	Gasoline blend discontinuity # 1 Sum. '10	# 2 Fall '10	# 3 Spr. '11	# 4 Fall '13	All four **pooled** together
A: Dependent variable: Mean (of demeaned) log PM2.5 concentration, by site-date					
CCT **E25, not E20, gasoline** (yes=1)			-0.235	0.119	0.028
i.e., higher ethanol content			(0.186)	(0.085)	(0.074)
Bandwidth			25	29	33
Number of observations			49	148	236
IK **E25, not E20, gasoline** (yes=1)			-0.077	0.057	0.018
i.e., higher ethanol content			(0.117)	(0.078)	(0.067)
Bandwidth			52	48	50
Number of observations			101	250	359
No. of sites (site-discontinuity pairs)	0	0	1	3	4
B: Dependent variable: Mean (of demeaned) log CO concentration, by site-date					
CCT **E25, not E20, gasoline** (yes=1)	-0.078	0.017	-0.083**	0.036	-0.018
i.e., higher ethanol content	(0.058)	(0.051)	(0.033)	(0.028)	(0.019)
Bandwidth	22	39	30	26	32
Number of observations	461	781	565	730	2,811
IK **E25, not E20, gasoline** (yes=1)	-0.015	0.017	-0.081**	0.037	-0.021
i.e., higher ethanol content	(0.026)	(0.051)	(0.033)	(0.028)	(0.019)
Bandwidth	90	39	31	27	30
Number of observations	1,964	781	583	759	2,628
No. of sites (site-discontinuity pairs)	12	11	10	15	48
C: Dependent variable: Mean (of demeaned) log NOx concentration, by site-date					
CCT **E25, not E20, gasoline** (yes=1)	-0.101	-0.121	-0.107*	0.004	-0.077**
i.e., higher ethanol content	(0.066)	(0.092)	(0.061)	(0.038)	(0.035)
Bandwidth	24	32	20	35	22
Number of observations	465	653	403	858	1,869
IK **E25, not E20, gasoline** (yes=1)	-0.023	-0.093	-0.023	-0.011	-0.040
i.e., higher ethanol content	(0.051)	(0.072)	(0.041)	(0.032)	(0.025)
Bandwidth	51	53	41	47	47
Number of observations	999	1,107	838	1,116	3,963
No. of sites (site-discontinuity pairs)	11	11	11	14	47

Notes: *** p<0.01, ** p<0.05, * p<0.1. An observation in the second step is a site-date pair within the bandwidth (in days from the cutoff point) specified by either the CCT or IK selection criterion, for all monitoring sites that are active in the Sao Paulo metropolitan area. See the text for the first-step procedure to demean hourly log pollutant concentrations (5pm to 8pm) using each 180-day sample, and to collapse residual concentrations to the site-date level taking the mean across hours. See the notes to Table H.3 for site-hour specific eighth-order polynomial trend, day-of-week fixed effects, meteorology (contemporaneous hour), thermal inversion (9am or 9pm reading, whichever is closest to the given hour), road traffic congestion (mean conditions 5pm to 8pm) and fuel price controls used to demean the data. Local linear regression estimates for 20 separate regressions. The remaining notes to Table H.4A apply here.

Table H.8: Local linear regression for Mean Ozone: Sensitivity to de-trending in the first step

2-5	Gasoline blend discontinuity # 1 Summer 2010	# 2 Fall 2010	# 3 Spring 2011	# 4 Fall 2013	All four **pooled** together
Dependent variable: Mean (of demeaned) log ozone concentration, by site-date					
CCT bandwidth criterion					
E25, not E20, gasoline (yes=1)					
8th-order polynomial (as	0.070**	0.112**	0.126***	0.078***	0.081***
in Tables H.4A and H.5)	(0.032)	(0.045)	(0.036)	(0.029)	(0.017)
9th-order polynomial	0.071**	0.114***	0.128***	0.078***	0.081***
	(0.032)	(0.044)	(0.037)	(0.029)	(0.017)
7th-order polynomial	0.077**	0.122***	0.126***	0.064**	0.083***
	(0.032)	(0.044)	(0.036)	(0.029)	(0.017)
6th-order polynomial	0.094***	0.126***	0.127***	0.045	0.081***
	(0.032)	(0.044)	(0.036)	(0.029)	(0.017)
No polynomial trend	0.100***	0.114**	0.118***	-0.036	0.052***
	(0.033)	(0.045)	(0.039)	(0.030)	(0.019)
IK bandwidth criterion					
E25, not E20, gasoline (yes=1)					
8th-order polynomial (as	0.066**	0.105***	0.118***	0.041	0.083***
in Tables H.4A and H.5)	(0.032)	(0.037)	(0.035)	(0.032)	(0.016)
9th-order polynomial	0.068**	0.104***	0.123***	0.040	0.085***
	(0.032)	(0.037)	(0.036)	(0.032)	(0.016)
7th-order polynomial	0.074**	0.135***	0.119***	0.035	0.086***
	(0.032)	(0.037)	(0.035)	(0.032)	(0.017)
6th-order polynomial	0.093***	0.133***	0.119***	0.019	0.084***
	(0.032)	(0.037)	(0.035)	(0.032)	(0.017)
No polynomial trend	0.100***	0.126***	0.112***	-0.053	0.055***
	(0.033)	(0.038)	(0.038)	(0.033)	(0.018)
No. of sites (site-disc. pairs)	12	12	12	17	53

Notes: *** p<0.01, ** p<0.05, * p<0.1. An observation in the second step is a site-date pair within the bandwidth specified by either the CCT or IK selection criterion, for all monitoring sites that are active in the Sao Paulo metropolitan area. Compared to the specification reported in Tables H.4A and H.5—which for convenience here we reproduce under "8th-order polynomial"—we vary the order of the (site-hour specific) polynomial trend that, in the first step, accounts for seasonal and unobservable trends over each 180-day sample. The bandwidth and (number of) observations are fixed at Table H.4A (or Table H.5) values (and for brevity are omitted), but estimates are similar if bandwidth is re-selected for each second-step sample. Local linear regression estimates for 50 separate regressions. Robust standard errors in parentheses.

Table H.9: Local linear regression for Mean Ozone: Bootstrap standard errors, clustered at site-date or site-week level

	Gasoline blend discontinuity			
2-5	# 1	# 2	# 3	# 4
	Summer '10	Fall '10	Spring '11	Fall '13
Dependent variable: Mean (of demeaned) log ozone concentration, by site-date				
CCT bandwidth criterion				
E25, not E20, gasoline (yes=1)	0.070**	0.112**	0.126***	0.078***
Robust standard errors (Table H.4A)	(0.032)	(0.045)	(0.036)	(0.029)
Site-date bootst. s. e. (bandw. per original)	(0.026)	(0.031)	(0.026)	(0.023)
Site-date bootst. s. e. (bandw. per bootstrap)	(0.026)	(0.033)	(0.032)	(0.021)
Site-week bootst. s. e. (bandw. per original)	(0.026)	(0.029)	(0.026)	(0.024)
Site-week bootst. s. e. (bandw. per bootstrap)	(0.025)	(0.032)	(0.034)	(0.027)
IK bandwidth criterion				
E25, not E20, gasoline (yes=1)	0.066**	0.105***	0.118***	0.041
Robust standard errors (Table H.4A)	(0.032)	(0.037)	(0.035)	(0.032)
Site-date bootst. s. e. (bandw. per original)	(0.025)	(0.025)	(0.026)	(0.024)
Site-date bootst. s. e. (bandw. per bootstrap)	(0.023)	(0.032)	(0.030)	(0.019)
Site-week bootst. s. e. (bandw. per original)	(0.025)	(0.023)	(0.025)	(0.023)
Site-week bootst. s. e. (bandw. per bootstrap)	(0.023)	(0.031)	(0.031)	(0.021)
Number of sites	12	12	12	17

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$ (p-values are based on the largest standard errors that are listed). Coefficients and robust standard errors estimated by local linear regression following Calonico et al. (2014a) as in Table H.4A, where a second-step observation is a site-date pair within the bandwidth specified by either the CCT or IK selection criterion, for all monitoring sites that are active in the Sao Paulo metropolitan area. To account for sampling variation in the first step that generates residuals $\tilde{y}_{iht}$ at the site-hour-date level, we repeat the two-step procedure on 400 bootstrap samples clustered at the site-date level or, alternatively, clustered at the site-week level. (Site-date or site-week) "bootst. s. e. (bandw. per original)" shows bootstrap standard errors (standard deviation over each set of 400 estimates) when the bandwidth is fixed at Table H.4A values, based on the original sample. "bootst. s. e. (bandw. per bootstrap)" shows bootstrap standard errors when the bandwidth is re-selected according to each bootstrap sample.

Table H.10: Local linear regression for Mean Ozone: Sensitivity to site, not site-hour, controls

	Gasoline blend discontinuity			
2-5	# 1	# 2	# 3	# 4
	Summer 2010	Fall 2010	Spring 2011	Fall 2013
Dependent variable: Mean (of demeaned) log ozone concentration, by site-date				
CCT bandwidth criterion				
E25, not E20, gasoline (yes=1)	0.100***	0.116**	0.060	0.100***
i.e., higher ethanol content	(0.035)	(0.045)	(0.038)	(0.034)
Bandwidth	29	24	26	27
Number of observations	619	528	574	851
IK bandwidth criterion				
E25, not E20, gasoline (yes=1)	0.093***	0.109***	0.058	0.083**
i.e., higher ethanol content	(0.034)	(0.038)	(0.035)	(0.035)
Bandwidth	31	33	31	25
Number of observations	663	733	693	786
Number of sites	12	12	12	17

Notes: *** p<0.01, ** p<0.05, * p<0.1. An observation in the second step is a site-date pair within the bandwidth (in days from the cutoff point) specified by either the CCT or IK selection criterion, for all monitoring sites that are active in the Sao Paulo metropolitan area. Compared to the specification reported in Table H.4A, we interact first-step controls only with site, rather than site-hour, fixed effects. Local linear regression estimates. Robust standard errors in parentheses.

Table H.11: OLS regression for Ozone using the 180-day sample: Sensitivity to linear phase-in

	Gasoline blend discontinuity			
2-5	# 1	# 2	# 3	# 4
	Summer 2010	Fall 2010	Spring 2011	Fall 2013
Dependent variable: Log ozone concentration, by site-hour-date				
Abrupt discontinuity (Table H.3)				
E25, not E20, gasoline (yes=1)	0.110**	0.157***	0.117***	0.116***
i.e., higher ethanol content	(0.043)	(0.050)	(0.036)	(0.039)
R^2	82%	86%	84%	83%
Linear phase-in over 3 days				
E25, not E20, gasoline (yes=1)	0.123**	0.167***	0.073*	0.192***
i.e., higher ethanol content	(0.047)	(0.055)	(0.042)	(0.048)
R^2	82%	86%	84%	84%
Linear phase-in over 5 days				
E25, not E20, gasoline (yes=1)	0.080	0.113**	0.082*	0.216***
i.e., higher ethanol content	(0.050)	(0.056)	(0.044)	(0.057)
R^2	82%	85%	84%	83%
Number of observations	9,852	10,056	10,122	13,487
Number of regressors	2,161	2,161	2,101	2,976
Number of sites	12	12	12	17

Notes: *** p<0.01, ** p<0.05, * p<0.1. An observation is a site-hour-date triple. The sample includes all hours between 12pm and 4pm on all dates that are plus and minus 90 days of the cutoff point. The dependent variable is the logarithm of hourly ozone concentration (μg/m^3). Compared to the specification reported in Table H.3, we consider a linear phase-in over three days (resp., five days) from the date cutoff specified in the sharp design, i.e., February 4 to 6 (resp., 4 to 8), 2010, May 4 to 6 (resp., 4 to 8), 2010, October 4 to 6 (resp., 4 to 8), 2011, or May 4 to 6 (resp., to 8), 2013, with weights on successive days increasing by 0.33 (resp., 0.2). For example, for a 3-day phase-in and the policy shifts from E20 to E25, weights are 0.33 on calendar day 4, 0.67 on day 5, and 1 on day 6 and beyond. OLS estimates for 12 separate regressions. Standard errors, in parentheses, are clustered by site-date pair.

Table H.12: Local linear regression for E20/E25 gasoline and E100 ethanol prices at the pump

	Gasoline blend discontinuity		
2-4	# 1, Summer '10	# 2, Fall '10	# 3, Spring '11
A: Dependent variable: Price of E20/E25 gasoline (BR$/liter)			
CCT bandwidth criterion & 2nd-step obs. is an individual station-date pair			
E25, not E20, gasoline (yes=1)	-0.008	-0.008	-0.011
i.e., higher ethanol content	(0.007)	(0.007)	(0.007)
Bandwidth	40	37	31
CCT bandwidth criterion & 2nd-step obs. is a zipcode-station brand-date triple			
E25, not E20, gasoline (yes=1)	-0.008	-0.005	-0.016*
i.e., higher ethanol content	(0.008)	(0.010)	(0.010)
Bandwidth	42	31	24
IK bandwidth criterion & 2nd-step obs. is an individual station-date pair			
E25, not E20, gasoline (yes=1)	-0.004	-0.007	-0.005
i.e., higher ethanol content	(0.005)	(0.006)	(0.005)
Bandwidth	81	52	52
IK bandwidth criterion & 2nd-step obs. is a zipcode-station brand-date triple			
E25, not E20, gasoline (yes=1)	-0.006	-0.004	-0.007
i.e., higher ethanol content	(0.006)	(0.007)	(0.006)
Bandwidth	60	53	69
B: Dependent variable: Price of E100 ethanol (BR$/liter)			
CCT bandwidth criterion & 2nd-step obs. is an individual station-date pair			
E25, not E20, gasoline (yes=1)	0.038	0.003	-0.003
i.e., higher ethanol content	(0.042)	(0.036)	(0.008)
Bandwidth	10	10	21
CCT bandwidth criterion & 2nd-step obs. is a zipcode-station brand-date triple			
E25, not E20, gasoline (yes=1)	-0.015	0.017	-0.006
i.e., higher ethanol content	(0.051)	(0.023)	(0.010)
Bandwidth	11	11	19
IK bandwidth criterion & 2nd-step obs. is an individual station-date pair			
E25, not E20, gasoline (yes=1)	0.008	-0.015	-0.003
i.e., higher ethanol content	(0.007)	(0.01)	(0.008)
Bandwidth	24	21	22
IK bandwidth criterion & 2nd-step obs. is a zipcode-station brand-date triple			
E25, not E20, gasoline (yes=1)	-0.003	-0.003	-0.004
i.e., higher ethanol content	(0.010)	(0.013)	(0.008)
Bandwidth	23	21	27

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. An observation in a second-step local linear regression is either an individual station by date pair (shown in the top half of panels A and B) or a 2-digit zipcode-station brand-date triple (bottom half of each panel), for all Sao Paulo city stations surveyed or zipcode-brand combinations within the bandwidth (in days from the cutoff point) specified by either the CCT or IK selection criterion. In a first-step demeaning procedure, we regress station-date level prices observed over a period of minus 90 to plus 90 days from a cutoff point on an eight-order polynomial trend, station brand fixed effects and 3-digit zipcode fixed effects. For the second-step estimates shown in the bottom half of each panel, we collapse residual prices resulting from the first step to the zipcode-brand-date level taking the mean across stations within each zipcode-brand combination. Local linear regression estimates for 24 separate regressions. Robust standard errors in parentheses.

Table H.13: Mean ozone levels (afternoon hours) and E100 ethanol shares in the bi-fuel fleet, or the entire light-vehicle fleet

Dependent variable:	Ozone	Ozone	Ozone	Log Ozone	Log Ozone	Log Ozone
	(1)	(2)	(3)	(4)	(5)	(6)
Specification (2nd step model): Share in the:	Bi-fuel fleet	Bi-fuel fleet	Light fleet	Bi-fuel fleet	Bi-fuel fleet	Light fleet
	Est.+Bootst.	Estimated+IV	Aggreg. Data	Est.+Bootst.	Estimated+IV	Aggreg. Data
In-sample variation in the E100 ethanol share:	24% to 89%	24% to 89%	18% to 49%	24% to 89%	24% to 89%	18% to 49%
In-sample variation in the E20/E25 gasoline share:	76% to 11%	76% to 11%	82% to 51%	76% to 11%	76% to 11%	82% to 51%
E100 ethanol share (i.e., 1 - E20/E25 gasoline share)	21.6***	26.6***	30.3***	0.37***	0.41***	0.77***
	(7.4)	(6.8)	(11.7)	(0.13)	(0.12)	(0.20)
Control variables						
Year-site fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Week-of-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-week fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Radiation (per 100 W/m^2 or log)	3.4***	3.4***	3.4***	0.33***	0.33***	0.33***
	(0.3)	(0.3)	(0.3)	(0.02)	(0.03)	(0.03)
Temperature (per 1^0C or log)	3.6***	3.6***	3.5***	1.36***	1.36***	1.36***
	(0.2)	(0.2)	(0.2)	(0.09)	(0.09)	(0.09)
Relative humidity (per 10 percent or log)	-5.3***	-5.3***	-5.4***	-0.48***	-0.47***	-0.48***
	(0.4)	(0.4)	(0.4)	(0.04)	(0.04)	(0.04)
Wind speed (per 1 m/s or log)	-10.8***	-10.8***	-10.7***	-0.13***	-0.13***	-0.12***
	(0.9)	(0.8)	(0.8)	(0.03)	(0.03)	(0.03)
Precipitation indicators (< 0.5, 0.5-2, > 2 mm/h)	Yes	Yes	Yes	Yes	Yes	Yes
Thermal inversion indicators (base at 0-199m, 200-499m)	Yes	Yes	Yes	Yes	Yes	Yes
Direction from which wind blows (site-specific by quadrant)	Yes	Yes	Yes	Yes	Yes	Yes
Morning road traffic congestion indicators (several)	Yes	Yes	Yes	Yes	Yes	Yes
Greater Sao Paulo beltway open (site-specific dummies)	Yes	Yes	Yes	Yes	Yes	Yes
Real price of diesel oil	Yes	Yes	Yes	Yes	Yes	Yes
R^2	75.9%	75.9%	75.9%	73.8%	73.8%	73.9%
Number of observations	18,768	18,768	18,768	18,767	18,767	18,767
Number of regressors	207	207	207	207	207	207
Mean value of dependent variable (μg/m^3 or log)	69.4	69.4	69.4	4.1	4.1	4.1

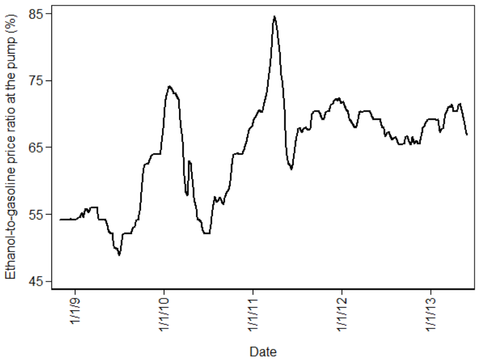
Notes: *** p<0.01, ** p<0.05, * p<0.1. An observation is a date-site pair. The sample includes all 12 monitoring sites (as in Salvo and Geiger 2014) but an extended period, from November 1, 2008 to May 31, 2013, including all weeks of the year and all days of the week. The dependent variable is the mean concentration (μg/m^3) between 12pm and 4pm on a given date and site (columns 1 to 3), or its logarithm (columns 4 to 6). Radiation, temperature, humidity, and wind speed in the recorded unit (columns 1 to 3), or its logarithm (columns 4 to 6). 2SLS estimates in columns 2 and 5, with the median ethanol-to-gasoline price ratio across retail stations instrumenting for the predicted E100 ethanol share in the bi-fuel fleet. OLS estimates in the remaining columns. In columns 3 and 6, the E100 ethanol share is calculated from reported aggregate distributor shipments that serve the entire state-level fleet of passenger cars and motorcycles. Standard errors are in parentheses. In columns 1 and 4, standard errors are calculated by bootstrapping (200 samples each): (i) the consumer-level fuel choice data, to account for sampling variation in the predicted E100 ethanol share, and (ii) the pollutant-meterology-traffic data, clustering by date. In the remaining columns, standard errors are clustered by date.

Appendix I

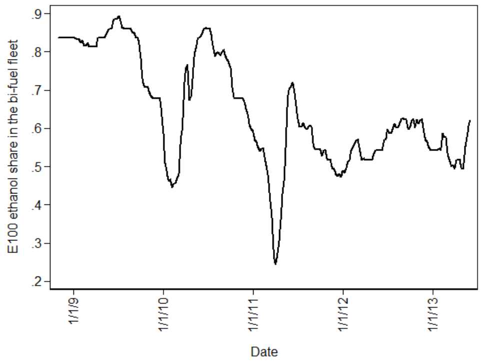
Chapter 3: Figures

Figure I.1: Fuel price ratio and usage

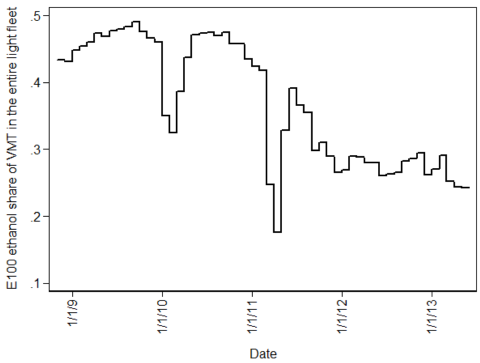
(a) Ethanol-to-gasoline price ratio



(b) E100 ethanol share in the bi-fuel fleet (choice probability)



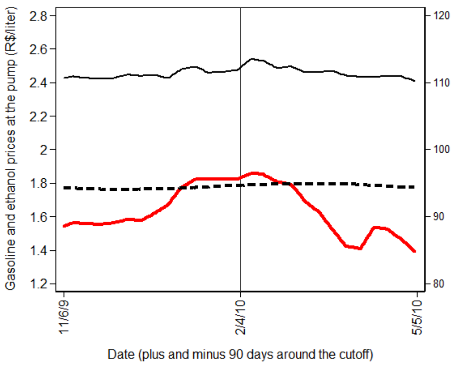
(c) E100 ethanol share of E20, E25 and E100 shipments (VMT equivalent share)



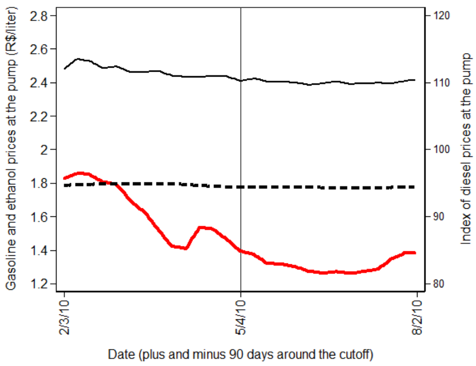
Notes: Evolution of: (a) the E100 ethanol to E20/E25 gasoline price ratio at the pump, (b) E100 ethanol share in the bi-fuel fleet, and (c) E100 ethanol share of fuel consumed by passenger cars and motorcycles. The series in panel (b) is imputed from an estimated discrete-choice demand model (Salvo and Huse 2013). The series in panels (a) and (c) are based on ANP data—see Table H.1 notes for further details on panel (c).

Figure I.2: Fuel price

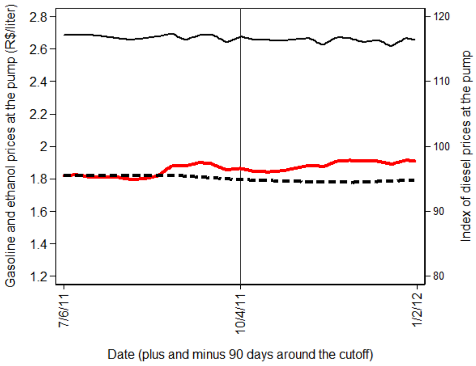
(a) Fuel prices at retail, discontinuity # 1



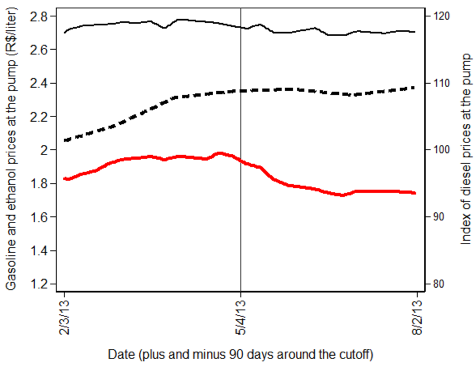
(b) Fuel prices at retail, discontinuity # 2



(c) Fuel prices at retail, discontinuity # 3

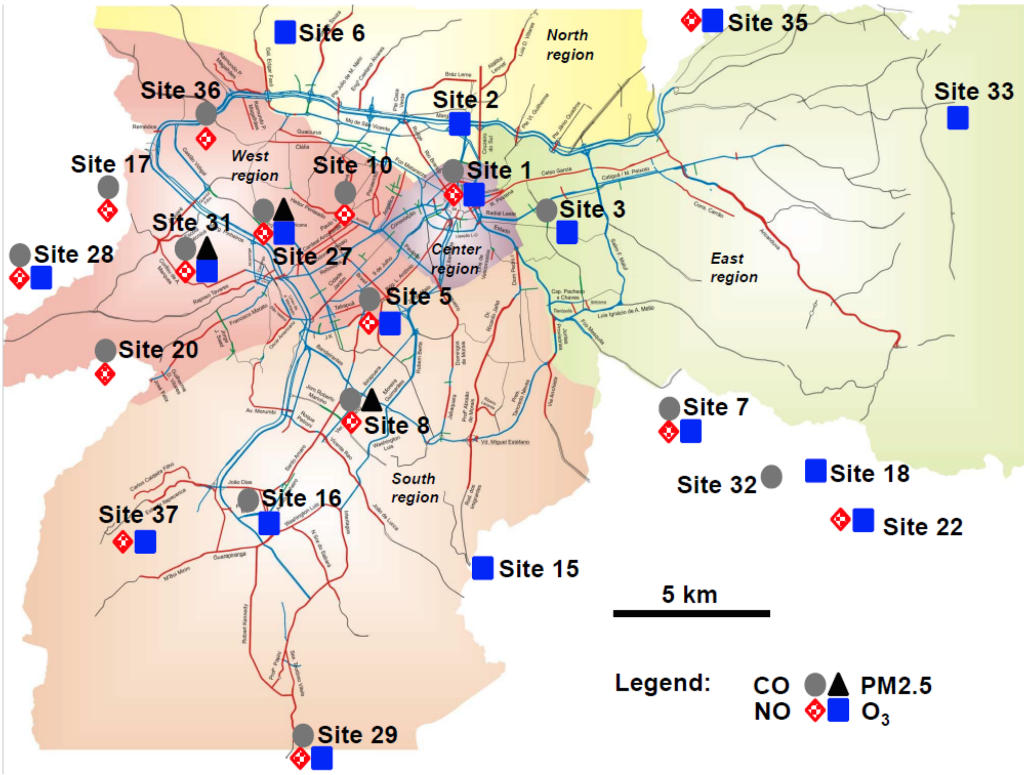


(d) Fuel prices at retail, discontinuity # 4



Notes: Variation in E20/E25 gasoline, E100 ethanol and diesel fuel prices at the pump around each discontinuity in the composition of gasoline (see Table H.1). A separate panel for each discontinuity marks the cutoff date with a vertical line and plots fuel price variation over the minus 90 to plus 90 days from this cutoff. E20/E25 gasoline (black, thin line) and E100 ethanol (red, thick line) prices in the left axis, in Brazilian Real $ per liter (means for regular-grade fuel—the dominant grade for both gasoline and ethanol—in a cross-section of about 350 retail stations surveyed weekly in the city of Sao Paulo, up to May 2013, and gasoline and ethanol price indices to August 2013). Diesel price index (dashed line) in the right axis (base October 30, 2008 = 100, for the Sao Paulo metropolitan area). Divide by 2 for a rough conversion from BR$ to US$ over the sample period. Source: ANP, IBGE IPCA.

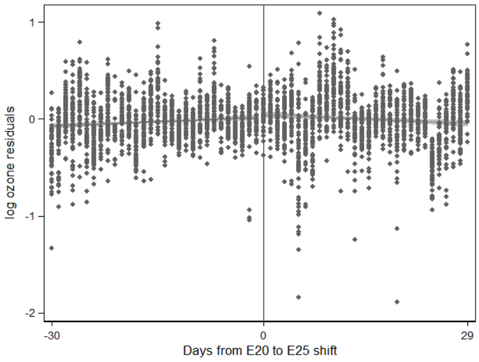
Figure I.3: Location of sites



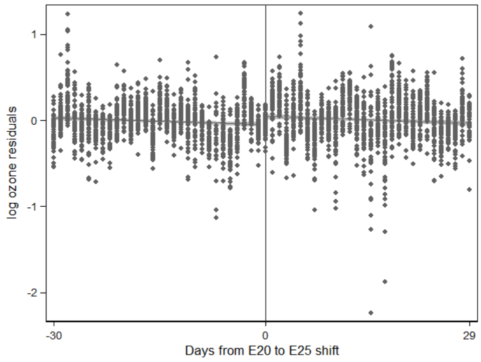
Notes: Location of the state EPA's air monitoring sites in the Sao Paulo metropolitan area (40 municipalities), monitoring ozone, PM2.5, CO and NOx (NO and NO_2), between 2010 and 2013, superimposed on the grid of road corridors that is monitored by the city's traffic authority for traffic congestion. Only regions in the city of Sao Paulo, a subset of the metropolitan area, are shaded. Air monitoring site names are: (1) Parque Dom Pedro II, (2) Santana, (3) Mooca, (5) Ibirapuera, (6) Nossa Senhora do O, (7) Sao Caetano do Sul, (8) Congonhas, (10) Cerqueira Cesar, (15) Diadema, (16) Santo Amaro, (17) Osasco, (18) Santo Andre-Capuava, (20) Taboao da Serra, (22) Maua, (27) Pinheiros, (29) Parelheiros, (31) IPEN-USP, (32) Santo Andre-Paco Municipal, (33) Itaim Paulista, (35) Guarulhos-Paco Municipal, (36) Marginal Tiete-Ponte dos Remedios, (37) Capao Redondo. Source: CETESB, CET.

Figure I.4: Residual plot

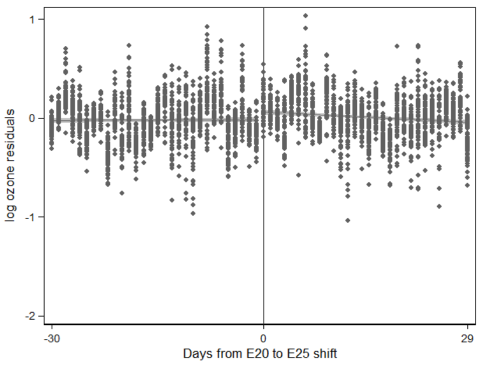
(a) Log ozone residuals, discontinuity # 1



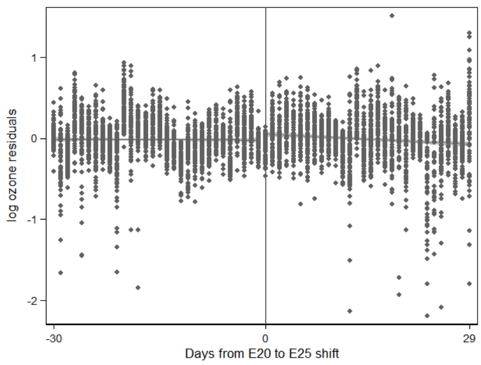
(b) Log ozone residuals, discontinuity # 2



(c) Log ozone residuals, discontinuity # 3



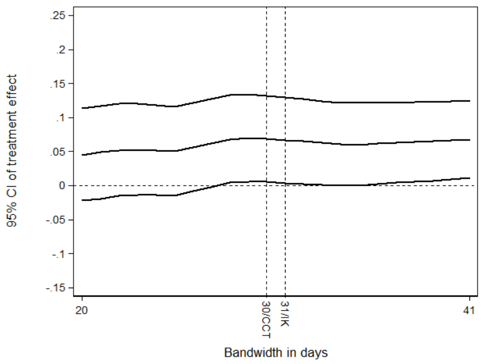
(d) Log ozone residuals, discontinuity # 4



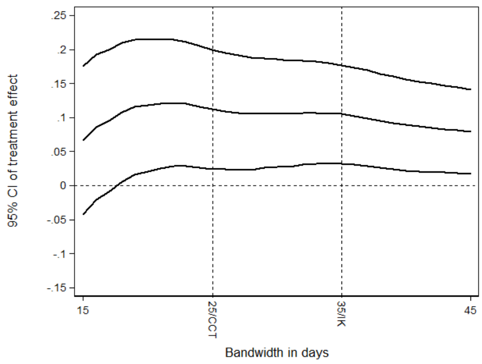
Notes: Demeaned log ozone concentrations, following the first step of the local linear regression estimation procedure, based on a 180-day sample separately by discontinuity. An observation is a site-hour-date triple, across afternoon hours (12pm to 4pm) and active ozone-monitoring sites. Only to illustrate, we plot fitted lines (with 95% confidence interval) on either side of each cutoff point fixing the bandwidth at 30 days. **Since the first and third discontinuity involved a shift, over time, from E25 to E20, the horizontal axis of panels (a) and (c) shows observations going backward in time. For example, in panel (a), the day labeled 0 is February 4, 2010, day 1 is February 3, 2010, and day -1 is February 5, 2010.**

Figure I.5: Robustness: bandwidth

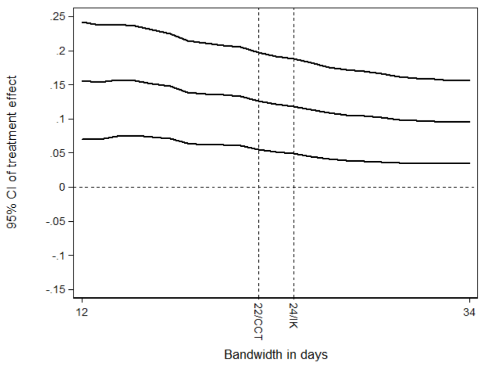
(a) E20 to E25 treatment effect, discontinuity # 1



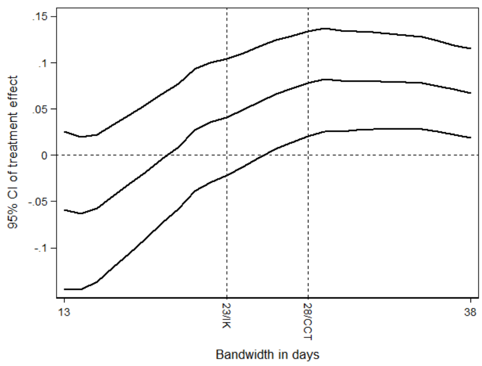
(b) E20 to E25 treatment effect, discontinuity # 2



(c) E20 to E25 treatment effect, discontinuity # 3



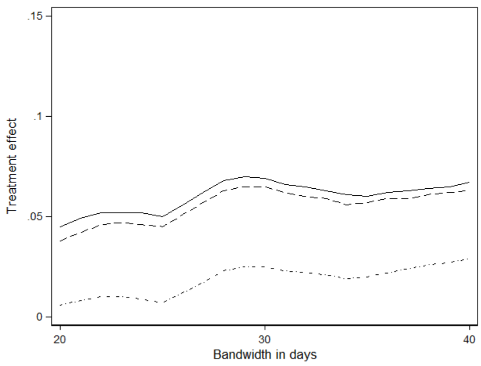
(d) E20 to E25 treatment effect, discontinuity # 4



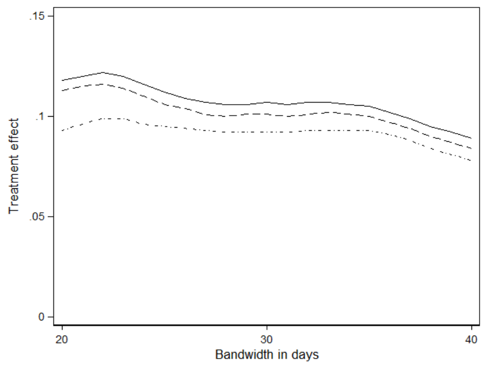
Notes: Local linear regression estimates of the E25 treatment effect on mean afternoon log ozone concentrations. 95% confidence intervals are plotted. Sensitivity analysis to variation in the optimal bandwidth, marked by dashed vertical lines (Table H.4A), according to the CCT and IK criteria.

Figure I.6: Robustness: demeaning method

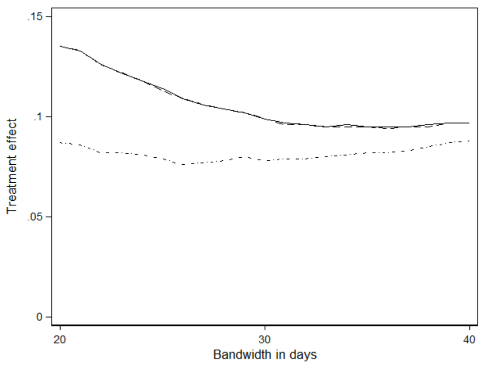
(a) E20 to E25 treatment effect, discontinuity # 1



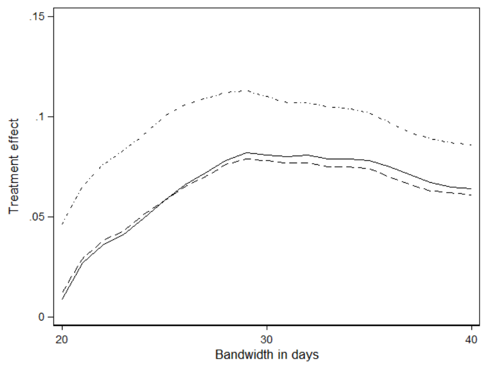
(b) E20 to E25 treatment effect, discontinuity # 2



(c) E20 to E25 treatment effect, discontinuity # 3



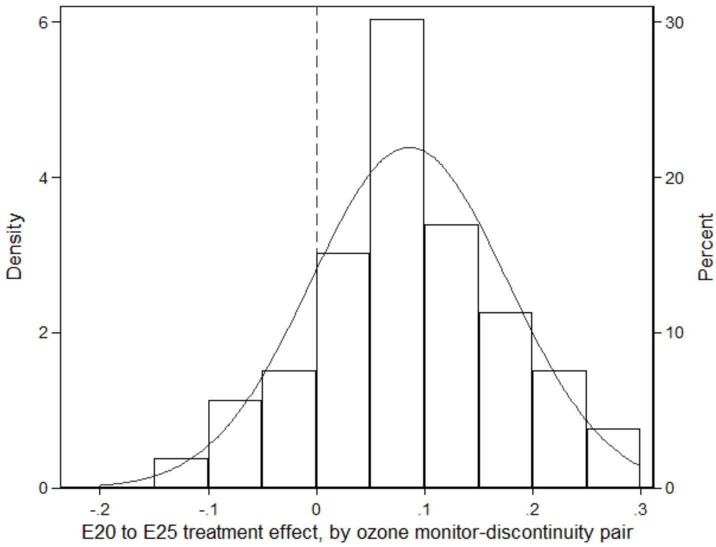
(d) E20 to E25 treatment effect, discontinuity # 4



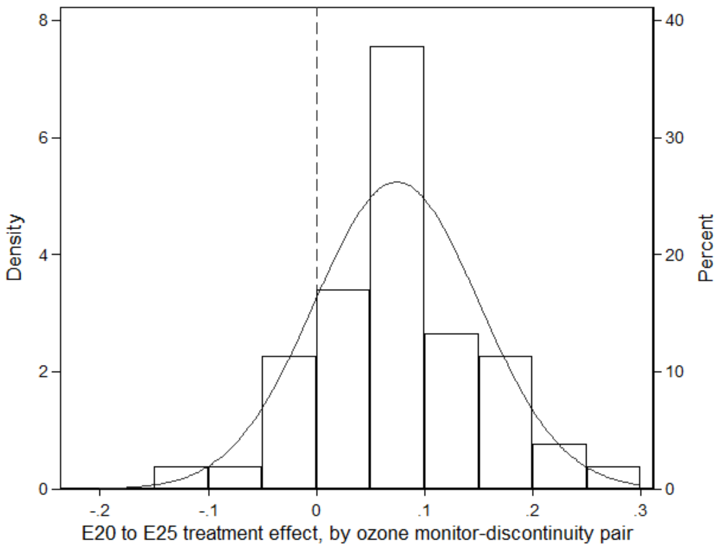
Notes: Local linear regression estimates of the E25 treatment effect on mean afternoon log ozone concentrations. Robustness to alternative procedures to demean the data in the first step, at varying bandwidths. Point estimates are plotted. Solid lines indicates preferred specification estimates (on means within site-date of demeaned log concentrations at the site-hour-date level). Dashed lines indicate estimates directly on demeaned log concentrations at the site-hour-date level. Dotted lines indicate estimates on demeaned log concentrations, where log concentrations are averages across hours within site-date.

Figure I.7: Effects of single sites

(a) CCT bandwidth criterion



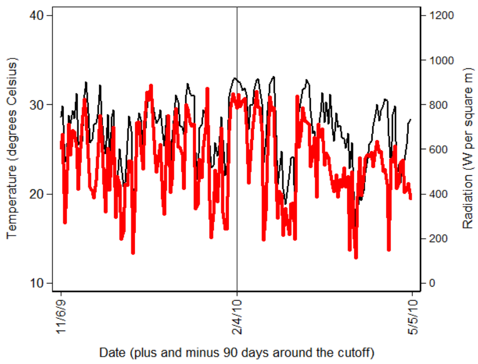
(b) IK bandwidth criterion



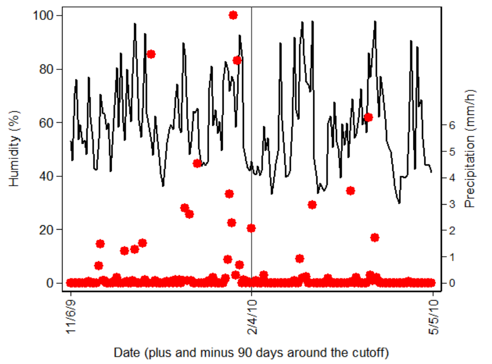
Notes: Local linear regression estimates of the E25 treatment effect on mean afternoon log ozone concentrations, implementing the second step separately by ozone monitoring site. As in our preferred specification, Table H.4A, local linear regression is implemented separately also by discontinuity. Each plot shows the distribution of 53 (i.e., $12 + 12 + 12 + 17$) estimated site by discontinuity treatment effects, $\hat{\alpha}_{id}$, where i and d index site and discontinuity, respectively.

Figure I.8: Meteorology and traffic: dis 1

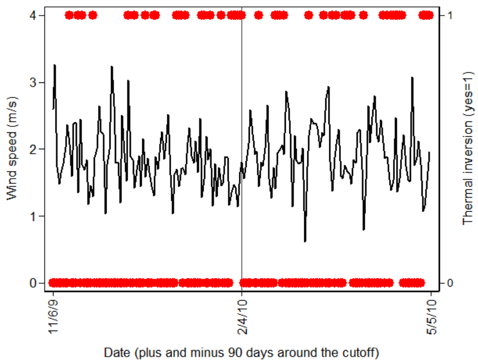
(a) Temperature (thin, left) & radiation (thick, right)



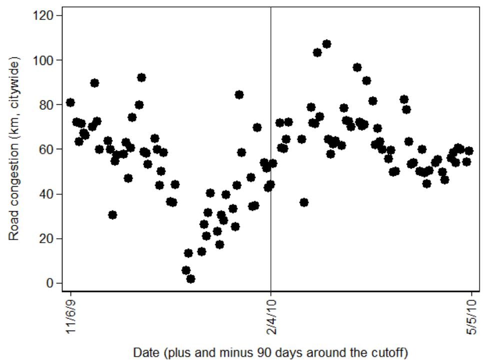
(b) Humidity (line, left) & precipitation (dots, right)



(c) Wind speed (line, left) & thermal inversion (dots, right)



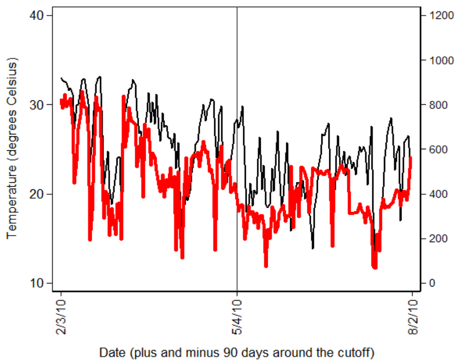
(d) Road congestion, weekday morning across city



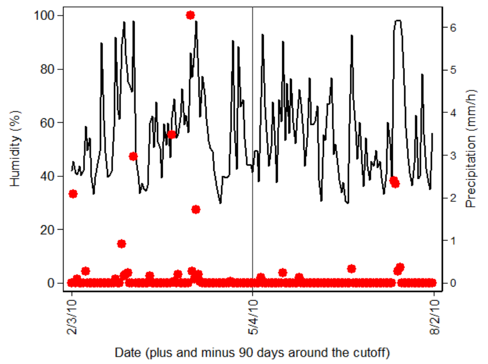
Notes: Variation in meteorological conditions and morning road congestion around discontinuity # 1. In each panel, the cutoff date is marked with a vertical line, and variation in meteorology or congestion is plotted over the minus 90 to plus 90 days from this cutoff: (a) temperature (black thin line, left axis, ^{0}C) and radiation (red thick line, right axis, W/m^2); (b) relative humidity (black thin line, left axis, %) and precipitation (red dots, right axis, mm/h); (c) wind speed (black thin line, left axis, m/s) and thermal inversion (red dots, right axis, yes=1); (d) road congestion in the morning across the city (km of extension). We plot the mean, by date, across hourly readings from 12pm to 4pm, except for thermal inversion where reading is at 9am, and road congestion for which we plot the mean from 7am to 11am on weekdays only.

Figure I.9: Meteorology and traffic: dis 2

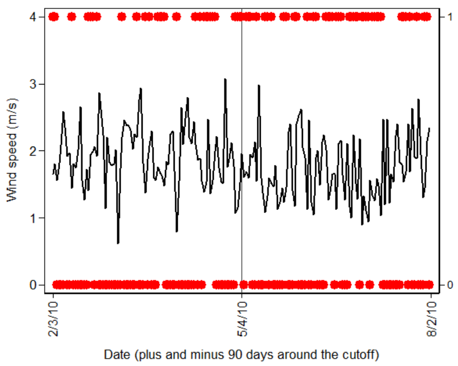
(a) Temperature (thin, left) & radiation (thick, right)



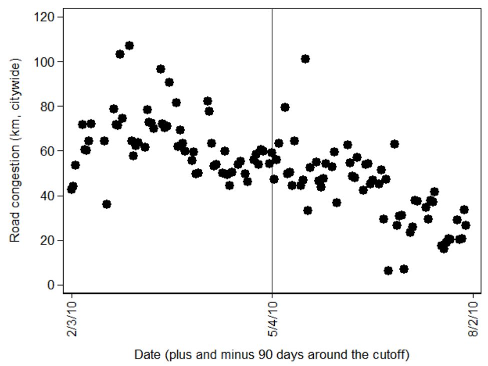
(b) Humidity (line, left) & precipitation (dots, right)



(c) Wind speed (line, left) & thermal inversion (dots, right)



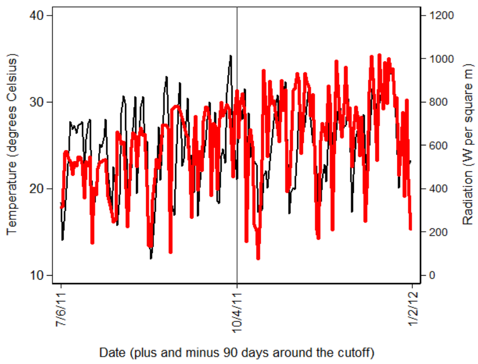
(d) Road congestion, weekday morning across city



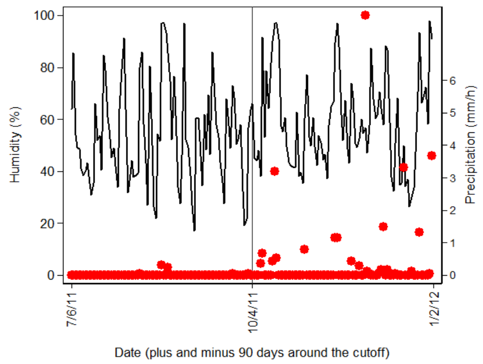
Notes: Variation in meteorological conditions and morning road congestion around discontinuity # 2. In each panel, the cutoff date is marked with a vertical line, and variation in meteorology or congestion is plotted over the minus 90 to plus 90 days from this cutoff: (a) temperature (black thin line, left axis, ^{0}C) and radiation (red thick line, right axis, W/m^2); (b) relative humidity (black thin line, left axis, %) and precipitation (red dots, right axis, mm/h); (c) wind speed (black thin line, left axis, m/s) and thermal inversion (red dots, right axis, yes=1); (d) road congestion in the morning across the city (km of extension). We plot the mean, by date, across hourly readings from 12pm to 4pm, except for thermal inversion where the reading is at 9am, and road congestion for which we plot the mean from 7am to 11am on weekdays only.

Figure I.10: Meteorology and traffic: dis 3

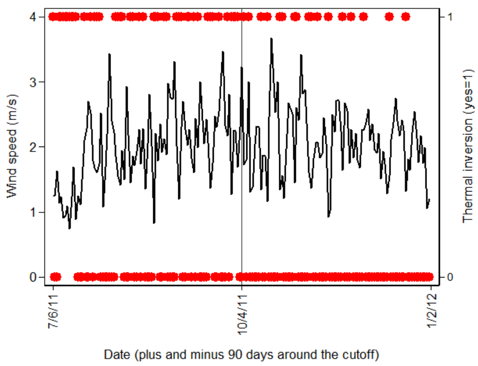
(a) Temperature (thin, left) & radiation (thick, right)



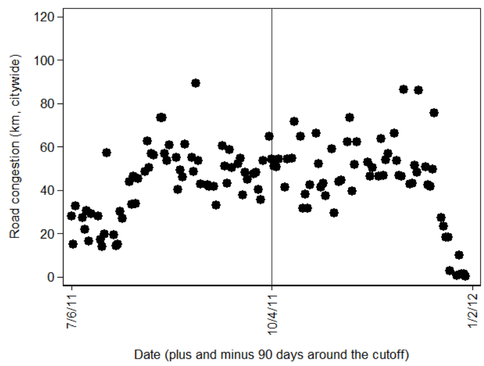
(b) Humidity (line, left) & precipitation (dots, right)



(c) Wind speed (line, left) & thermal inversion (dots, right)



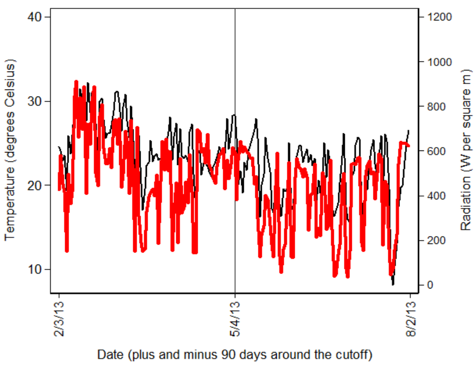
(d) Road congestion, weekday morning across city



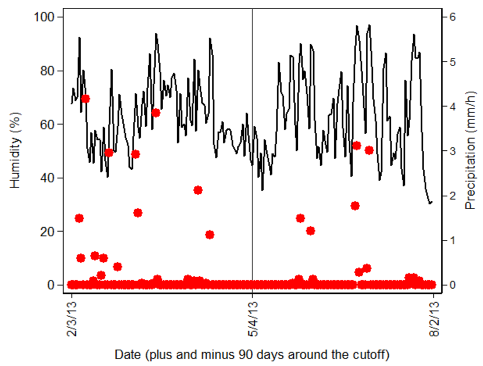
Notes: Variation in meteorological conditions and morning road congestion around discontinuity # 2. In each panel, the cutoff date is marked with a vertical line, and variation in meteorology or congestion is plotted over the minus 90 to plus 90 days from this cutoff: (a) temperature (black thin line, left axis, ^{0}C) and radiation (red thick line, right axis, W/m^2); (b) relative humidity (black thin line, left axis, %) and precipitation (red dots, right axis, mm/h); (c) wind speed (black thin line, left axis, m/s) and thermal inversion (red dots, right axis, yes=1); (d) road congestion in the morning across the city (km of extension). We plot the mean, by date, across hourly readings from 12pm to 4pm, except for thermal inversion where the reading is at 9am, and road congestion for which we plot the mean from 7am to 11am on weekdays only.

Figure I.11: Meteorology and traffic: dis 4

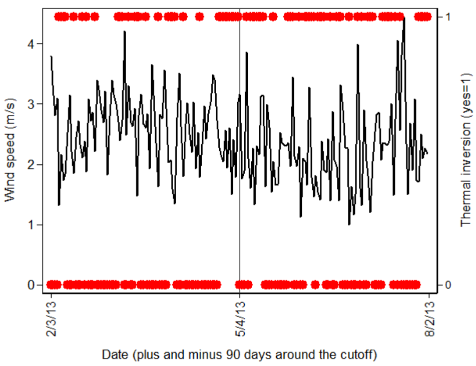
(a) Temperature (thin, left) & radiation (thick, right)



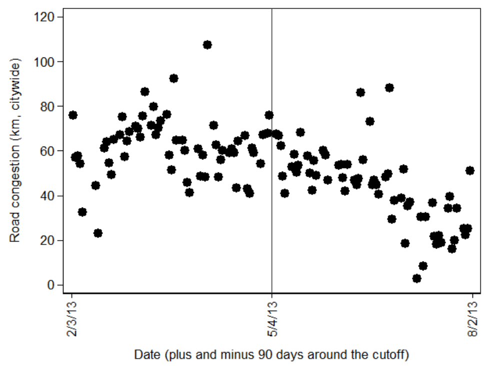
(b) Humidity (line, left) & precipitation (dots, right)



(c) Wind speed (line, left) & thermal inversion (dots, right)



(d) Road congestion, weekday morning across city



Notes: Variation in meteorological conditions and morning road congestion around discontinuity # 2. In each panel, the cutoff date is marked with a vertical line, and variation in meteorology or congestion is plotted over the minus 90 to plus 90 days from this cutoff: (a) temperature (black thin line, left axis, ^{0}C) and radiation (red thick line, right axis, W/m^2); (b) relative humidity (black thin line, left axis, %) and precipitation (red dots, right axis, mm/h); (c) wind speed (black thin line, left axis, m/s) and thermal inversion (red dots, right axis, yes=1); (d) road congestion in the morning across the city (km of extension). We plot the mean, by date, across hourly readings from 12pm to 4pm, except for thermal inversion where the reading is at 9am, and road congestion for which we plot the mean from 7am to 11am on weekdays only.