Macalester Journal of Economics

Volume 30 Spring 2020

Table of Contents

Published annually by the Macalester College Department of Economics and Omicron Delta Epsilon

Macalester Journal of Economics

Volume 30 *Spring 2020*

Contents

FOREWORD

The Macalester College chapter of Omicron Delta Epsilon, the international honors society in economics, proudly edits the Macalester Journal of Economics every year. (Yes, even *this* year.) The editors—Christopher Werbos '20 (Arlington, VA), Vincent Mougin '20 (Seattle, WA), Camille Baker '20 (Chicago, IL), Yannick Laurent '20 (Los Altos, CA), Natasha Mwonga '20 (Nairobi, Kenya), William Sandy '20 (Shorewood, WI), and Esther Swehla '20 (Los Angeles, CA)—have carefully selected five papers on a variety of important topics. These papers are a sample of the research that our students produced in the last academic year.

The first two papers in the journal ask important questions in the field of economic growth and development. Jasmine Davidson '20 (Redwood City, CA) asks: is time spent playing well spent? Does more playtime result in any measurable benefit for the kids involved? In an era of hyper-competitive parenting, where children lives are oftentimes scheduled to the minute, it's worth asking whether playing a bit (in an unorganized way) provides some long-run advantage. Jasmine looks at Peruvian data and finds mixed evidence on the effect of play over child development: while there is no broad conclusion to be reached, children in the lower income deciles seem to benefit from playtime later in their lives, as measured by academic achievement tests. In this sense, playing might be beneficial to some kids, but remains fun for every one of them.

Still in the orbit of growth and development, a perennial topic concerns the determinants of growth over the very long run. This is especially relevant for countries that received a colonization shock early in their history: why did the path of development of, say, the United States, is dramatically different from that of Mexico? One of the most plausible explanations involves institutions: broadly speaking, countries where the institutional equilibrium called for an extractive nature (Mexico) fared worse than those where the equilibrium led to a social contract that resembled the institutions found in the colonizing country (United States). But these explanations condense hundreds of years of history across many countries into a couple of paragraphs. So, how exactly is an extractive state formed? This is what the paper by Paul Cosme '22 (Quezon City, Philippines) aims to answer. He provides an historical account of the Philippines and shows us how an extractive regime arises. While many observers argue that the Marcos regime was to blame, Paul shows that the Marcos regime is instead the conclusion of a very long process that started (not surprisingly) in the Spanish colonial period.

From growth and development we transition to urban economics. It is well known that “good” houses are pricey, but what determines whether a house is “good” to begin with? Sam Jakshtis ’21 (Hardwick, MA) sets to identify the determinants of home desirability within the Twin Cities. Using a hedonic regression model and data from 24 ZIP codes in Minneapolis and Saint Paul, he looks at how single-family home characteristics determine the purchase price. Somewhat puzzling, he finds that an increase in the walk score—how close the home is to a central business district—reduces the price of a single-family home. Other interesting results include a non-linear relationship between price and the home’s age (so in this sense, a house is like wine: older is sometimes better).

From urban economics we now switch to international trade. One of the main issues (if not *the* main issue) on the radar of trade economists is the U.S.–China trade war. Notwithstanding the back-and-forth of who pays the tariffs in the first place (I shall not go there because we all know the answer), it is reasonable to ask whether China is looking for a way to go around these measures; for example, China has relied on transshipment—rerouting trade through other countries—as a strategy to avoid tariffs. Yannick Laurent ’20 (Los Altos, CA) asks whether China is once again relying on transshipment in the current escalation of tariffs: he finds evidence that the answer is on the affirmative. He estimates that roughly $800 million of exports from China are sent via transshipment; that said, while this seems like a large amount, it barely represents 0.15% of American imports from China in 2018!

The last paper takes us to the field of labor economics and concerns the minimum wage. Many journal pages have been spent discussing the effects of increases in this policy variable, especially over low-income families. But such discussions rarely discuss the intertemporal dimension of such changes: how do increases in the minimum wage affect low-income families, not only on impact, but also over time? Are they more likely to exit poverty in the medium run? Camille Baker ’20 (Chicago, IL) tackles this question using a family’s eligibility and participation in the Supplemental Nutrition Assistance Program (SNAP) as a proxy for its income. She finds that increases in the minimum wage reduce SNAP participation, but that these effects vanish within a year and a half. These findings should be interesting to economists and policymakers alike.

On behalf of my colleagues in the Economics Department, I am delighted to present the research of these talented students. I am confident that you will find it enlightening and be impressed by the value of a liberal arts education.

Mario Solis-Garcia
Associate Professor of Economics

Time to Play: The Relationship Between Time Spent Playing and Educational Outcomes in Peru

Jasmine Davidson

June 2020

Abstract

Every day, children around the world are playing. There has been plenty of research on the importance of different kinds of play, but very little on the importance of the quantity of play. Understanding the relationship between educational outcomes and the amount of time spent playing would allow parents to better structure their children's time and would settle the debate between psychologists and economists on whether play has inherent value for a child's future outcomes. I focus on Peru because conducting this research in a developing country context broadens the current research mostly focused on high-income countries. Using child-level, longitudinal data from the Young Lives Survey in Peru, I perform several regressions to better understand how time spent playing at age five is related to test scores and grade level at age fifteen. Ultimately, I find little evidence for a strong relationship, either positive or negative. However, I do find that more play is related to better math scores for children in the lowest wealth quartile, and lower educational attainment for children in the second-lowest wealth quartile. This suggests that a relationship between the quantity of play and educational outcomes may exist, but only for particular populations. Further study is needed to carefully untangle these relationships and settle this debate.

Keywords: Play, education, Peru

Acknowledgements

I would first like to acknowledge Amy Damon for her invaluable feedback and support throughout this project. I also want to thank Samantha Çakır and Tina Kruse for their comments on earlier versions of this work. The '19-'20 Macalester Economics Honors Seminar students also provided insightful feedback when I first began this project, for which I am immensely grateful. Finally, I would like to thank the Young Lives survey team for allowing me access to the data. Any errors in this project are solely my own.

1 Introduction

Economics and developmental psychology are at odds when it comes to the impact of the quantity of play on life outcomes. On the one hand, economists view play time essentially as a good (Becker, 1965), with no positive effects on a child's life outcomes. On the other hand, it is widely accepted among psychologists and neurobiologists, and even economists, that early experiences have a strong influence on the development of cognitive and social skills and on brain architecture and neurochemistry (Knudsen et al., 2006; Shonkoff et al., 2000). Developmental psychologists go one step further, viewing play as critically important in child development (Piaget, 1962). Indeed, there is already some evidence that play is important for cognitive development (Bergen, 2018; Nicolopoulou, 2018; Tamis-LeMonda et al., 2004; Urke et al., 2018), and even for longer-term outcomes (Schweinhart & Weikart, 1997; Gertler et al., 2014). Yet so far, there has been minimal research on the relationship between time spent playing and educational outcomes, and not in a developing country context. Thus, my research will focus on the question: "How does time spent playing correspond to later test scores and educational persistence for children in Peru?"

On the international stage, Finland is often looked to as a "gold standard" in education, and play is an important part of their model. Finland's education system is unregimented – short school days, low amounts of homework, and no national exams apart from the final one at the very end of high school (Finnish National Agency for Education, 2018). Especially in pre-primary education and early elementary school, play is prioritized as a method of learning (Walker, 2015; Hancock, 2011). Ever since the first Programme for International Student Assessment (PISA) exam was administered in 2000, Finland has scored highly compared to other OECD countries, so many have attempted to replicate its system in order to replicate its success. However, its scores have been steadily declining since it first made headlines in 2000, and the gap between rich and poor pupils has been steadily rising, so perhaps it has not been as successful as it might first appear (The Economist, 2019). Still, Finland does score quite highly, and perhaps its system allows students a better overall quality of life not reflected in PISA scores. In any case, Finland's relative wealth and homogeneous culture make it difficult to directly compare it and its play-based education strategies to Peru, so we need to examine Peru specifically to draw any conclusions about the play-educational outcomes relationship there.

1.1 Education in Peru

Peru is a developing country in South America with an average yearly per-capita income of about 6,530 USD (Worlddata.info, 2019). As of 2018, 81.9% of all adults over the age of 25 have at least completed primary school, 58.0% of all adults over 25 have completed upper secondary school, and these numbers seem to be trending steadily upwards (World Bank, 2018a,b). It is in this context that children go to school to learn and to improve their economic futures. There is one year of compulsory pre-school education (*educatión inicial*) that begins approximately at age five. There are then six years of primary school and five years of secondary school, with students completing school around age 17 (Clark, 2015).

Grade repetition has been decreasing in recent years; in 2007, about seven percent of students repeated a grade, while in 2017 only about three percent did so (UNESCO Institute for Statistics, 2019). Results are different for indigenous children, however; in 2007, schools with an indigenous student population over 50% had rates of grade repetition around 13%

(Cueto et al., 2009). Drop-out rates also differ by indigeneity; in 2007, the out-of-school rate was about eight percent, compared to the national average of about five percent (Cueto et al., 2010; UNESCO Institute for Statistics, 2019).

Learning outcomes, as measured by the 2018 PISA exam scores, are different along gendered and socio-economic lines (OECD, 2019). Girls outperformed boys in reading, but boys outperformed girls in math and science. Socio-economically advantaged students outperformed disadvantaged students in reading. While these are the only factors that PISA measured, they are certainly not the only ones impacting educational outcomes – another set of important factors are related to the child's early childhood environment.

1.2 Early Childhood Environment and Outcomes

There is a substantial body of literature on how a child's early environment impacts their development and future life outcomes. Glewwe et al. (2001) found that Filipino children with better nourishment, as measured by height-for-age, performed better in school, both because they entered school earlier and because of greater learning productivity per year of schooling. Gould et al. (2011) examined the 1949 Magic Carpet Operation, a natural experiment where over 50,000 Yemenite families were airlifted to Israel and scattered across the country essentially randomly. When in an early childhood environment with better sanitary and infrastructure conditions, these immigrant children were more likely to obtain higher education, marry at an older age, have fewer children, and work at age 55. The Moving to Opportunity project, which randomly assigned housing vouchers to low-income families and examined how moving when young impacts life outcomes, presents contrasting evidence. Children who moved to lower-poverty neighborhoods did not have significantly different educational outcomes (Sanbonmatsu et al., 2006) or physical health outcomes (Ludwig et al., 2013); the only statistically significant effect was that females who moved were less likely to have mental health problems (Ludwig et al., 2013).

Preschool can also be an important part of a child's early environment. Head Start is a public preschool program in the United States for disadvantaged children. Garces et al. (2002) found that participation in Head Start was associated with a significantly increased probability of completing high school and attending college for white children, and with a significantly lower likelihood of having been charged or convicted of a crime for black children. Similarly, the Perry Preschool Project was a two-year high-quality play-based preschool program with weekly home visits for black children living in poverty in Ypsilanti, Michigan. It led to greater educational and economic successes for the children and significantly reduced their crime rate (Schweinhart and Weikart, 1997), and the annual social rate of return on this preschool investment is estimated between seven and ten percent (Heckman et al., 2010). It is important to note that since both of these preschool programs were targeted towards specific demographics, their results may not be generalizable to all students.

In their cross-disciplinary examination of research in economics, developmental psychology, and neurobiology, Knudsen et al. (2006) corroborated the general findings above. They showed that early experiences are influential for the development of cognitive and social skills and on brain architecture and neurochemistry, and that it becomes more difficult for human skill development and neural circuitry to change over time. Thus, there is substantial evidence suggesting the importance of a child's early environment on their long-term outcomes. It is possible that an early environment conducive to lots of play may impact long-term outcomes, too. I will next examine how various disciplines have treated play, then discuss the current literature on the beneficial effects of play.

1.3 Economic Theory and Play

Economic theory has thus far rarely considered play and leisure to have any inherent value or potential to create better outcomes. Instead, leisure is generally considered in terms of its opportunity cost – like the cost of studying or working less (Becker, 1965; Gershuny, 2009; Sevilla et al., 2012). Thus, the theory goes, more play would lead to worse educational outcomes, worse test scores, and worse labor market outcomes.

Crispin and Kofoed (2019) offer evidence of the opportunity cost of leisure. Using the American Time Use Survey, they found that if a high school student was working, they were significantly less likely to participate in extracurricular activities, and they spent significantly fewer hours in extracurricular activities if they did participate. They also found that the fathers' education level was an important factor, and that while low-income students were less likely to participate in extracurriculars, this effect was largely driven by selection of students into the labor market. Pike et al. (2009) examined the opportunity cost of working in relation to grades, using the 2004 National Survey of Student Engagement (NSSE) to find that first-year university students in the United States working over 20 hours per week achieved lower grades. Both Pike et al. and Crispin and Kofoed present strong evidence of the existence of important trade-offs – working more means enjoying less leisure and lower academic achievement.

There is also evidence that studying increases academic achievement. While Schuman et al. (1985) found that there is at best only a small relationship between studying and grades for U.S. university students, Michaels and Miethe (1989) claimed that specification errors (the functional form and the omission of important variables) render these results spurious. Rather, Michaels and Miethe used their own survey for university students and found significant main and interactive effects of academic effort and college grades. More recently, Andrietti and Velasco (2015) found that there was a statistically significant and positive relationship between hours studying and grades for university students in Spain. Both of these studies imply an important trade-off: more leisure time, if it means less time spent studying, likely leads to lower grades.

1.4 Psychological Theory and Play

Psychologists and many educators approach play differently – they see it possible (and even likely) that play in and of itself has the potential to improve child development, and thus childhood outcomes. Maria Montessori and her philosophy of education famously emphasized a blend of freedom and structure using interactive teacher lessons, freely chosen activities, and engagement with peers – all of which were intrinsically rather than extrinsically rewarded (Lillard, 2013). While Montessori herself did not see the value in some types of play, her approach has become somewhat synonymous with playful learning.

But what exactly is play? Defining play is notoriously difficult, but it can be conceived along a few important dimensions (Pellegrini, 2009). The first is the "structural" dimension of play, which relates to the directly observable actions while a child is at play, like exaggerated movements, running, jumping, alternating roles, and "play face". The second is the functional dimension of play, which relates to how the child's actions resemble a functional behavior but do not serve that purpose – for example, using a play kitchen and pretending to make eggs (but not actually cooking anything). The final dimension is the causal dimension of play, which examines the contexts in which play is observed – it is interrupted by more serious concerns, is voluntary, and is characteristic of juveniles.

There are two foundational theories of play: that of Piaget, and that of Vygotsky. Piaget, a central figure in western theories of child development, essentially viewed play as a way for children to subordinate the world they encounter to their own points of view (assimilation), which counterbalances the forces of accommodation (the domination of the internal by the external) (Pellegrini, 2009). Vygotsky, another prominent child development theorist based out of the former USSR, viewed play as a form of reconciliation between wishes that cannot be fulfilled in reality (but can in fantasy) and societal norms limiting those choices, and this reconciliation happens by creating an imaginary world that also conforms to societal reality at some level (Pellegrini, 2009).

Empirical research on play has largely been based on these two theorists and can be categorized into four main types of play. Based on Piaget's (1962) theory of development, different types of play correspond to different stages in child development (Pellegrini, 2009). The earliest form of play is functional or practice play, and corresponds to the sensorimotor period of development, involving repetition of behaviors and routines. The next form of play is symbolic play, often known as pretend, fantasy, or make-believe play, and is "assimilative behavior where a behavior is taken out of its functional context", like using a banana to represent a telephone (Pellegrini, 2009, p.15). Another form of play is games-with-rules, where children play by rules that are defined *a priori*. A final category of play based more on Vygotsky's work is social play, either with adults or with peers.

1.5 Play and Child Development

Certain types of play may have a strong impact on child development. Bergen (2018) found in her review of neuroscience and psychology literature that sensorimotor play is an important precursor to other levels of child development. In a review of the literature, Nicolopoulou (2018) found that pretend play is beneficial to child development, although Lillard et al. (2013) did a similar review of the literature and found that existing evidence does not support causal claims of the unique importance of pretend play.

The relationship between social play and child development has been well-studied across cultures. Urke et al. (2018) used Demographic and Health Surveys (DHS) in Honduras and found that maternal psychosocial stimulation was significantly associated with early childhood development. Abimpaye et al. (2019) performed a randomized controlled trial of a parenting intervention on playful learning in Rwanda and found that children whose parents received the intervention had significantly higher child development scores. Tamis-LeMonda et al. (2004) used observational data from the National Head Start Evaluation Study in the United States and found that supportive parenting during play improved children's cognitive and language outcomes. Finally, Gertler et al. (2014) used a randomized controlled trial and found that facilitating mother-child play for stunted Jamaican toddlers developed their cognitive and socioemotional skills.

1.6 Play and Long-Term Outcomes

Play, and the improved cognitive development that follows, influences lifetime outcomes. As mentioned earlier, the Perry Preschool Project was a play-based preschool program that led to improved educational and economic outcomes alongside a significantly reduced crime rate (Schweinhart & Weikart, 1997). However, play was not necessarily the reason for these improved outcomes – the program also included weekly home visits and fostered strong parent-child relationships that lasted long after the child began formal schooling.

Gertler et al. (2014) were able to isolate the effects of play through a randomized controlled trial. They found that facilitating mother-child play for stunted Jamaican toddlers improved their cognitive and socioemotional skills, and that the intervention increased the children's earnings 20 years later by 25%, allowing them to catch up to their non-stunted peers. This improvement was not due to long-term increased maternal activities with the child – by age 7, there was no difference in maternal stimulation between the treatment and control groups. The improved incomes may have been due to increased parental investment in the child, however; as children exited the intervention period with higher skills, the parents may have recognized that schooling had higher returns than they previously realized, and by age 22, the treatment group had significantly more years of schooling attainment and a significantly higher proportion still enrolled in school. The improved incomes may also be due to the program improving children's skills enough so that families moved abroad to take advantage of better education and labor market opportunities; the migration rate of the treatment group was significantly higher than that of the control group. Still, both the Perry Preschool Project and the Gertler et al. study offer evidence that play may be important in determining future life outcomes.

1.7 Important Gaps

While there has been significant research on the importance of different types of play for child development and life outcomes, there has been very little work on the importance of time spent playing. Is simply allotting time for play sufficient to attain the beneficial outcomes related to play, or must the play be of a specific type? Pellegrini et al. (1995) lend support to the hypothesis that playing more positively impacts educational achievement – they found that inattention before recess was much higher than inattention after recess. But there have been no studies, to my knowledge, that directly examine the relationship between time spent playing and educational outcomes.

Much of the research on the importance of play for child development and life outcomes has been conducted in developed countries. With the exception of a number of studies on the importance of parent-child interaction (Gertler et al., 2014; Urke et al., 2018; Abimpaye et al., 2019), there has been minimal research on the importance of play in developing countries in general, and very little in Peru in particular. By leaving out developing countries, we risk believing a particular relationship to be universally true, when in fact the relationship may vary by country for cultural, socio-economic, or other reasons. It is not necessarily true that the play-education relationship is different in Peru than it is in the United States and other high-income countries – but given the opportunity to investigate this relationship in Peru, it makes sense to do so.

Thus, this paper will contribute to the literature in two ways: 1) by examining the effect of time spent playing, and 2) by conducting the analysis in a developing country context, specifically Peru. Based on the work of Gertler et al. (2014) and others, I hypothesize that children who spend more time playing have better educational outcomes than children who play less.

This paper proceeds as follows. In the next section, I outline the theoretical framework for my analysis, drawing on principles of utility maximization and the production function for academic achievement. Next, I discuss my empirical model and potential bias in my estimates. I then describe the Young Lives Survey data and present my results, finding little evidence for a strong relationship between quantity of play and educational outcomes in general, but some evidence that it may be important for specific groups. I conclude with a call for more studies on this relationship, and a warning to policymakers to be aware of the strong effects of wealth on educational outcomes.

2 Theoretical Framework

Children are usually not in charge of how they allocate their time – this is dictated by their parents/caregivers. Parents want to maximize their utility with respect to their child by choosing the amount of leisure enjoyed and income earned by the child. We can imagine this takes a Cobb-Douglass form:

$$U = C^{\alpha} L^{\beta} \tag{1}$$

where U is the parents' utility with respect to the child, C is consumption, L is leisure, and the parent is choosing these values for their child. The parents' utility maximization regarding the child is subject to a budget constraint:

$$Y = pC + wL \tag{2}$$

where Y is income, p is the price of goods, and w is the wage rate. This wage rate is not necessarily money that the child is earning directly, but rather can represent the wage someone else in the household is able to earn because the child is performing household tasks. The objective function is thus

$$(C, L, \lambda) = C^{\alpha} L^{\beta} - \lambda[Y - pC - wL] \tag{3}$$

with first order conditions of

$$\frac{\partial F}{\partial C} = \alpha C^{\alpha-1} L^{\beta} - \lambda p^{*} = 0 \tag{4}$$

$$\frac{\partial F}{\partial L} = \beta C^{\alpha} L^{\beta-1} - \lambda^{*} w = 0 \tag{5}$$

$$\frac{\partial F}{\partial \lambda} = -Y + pC + wL = 0 \tag{6}$$

Solving the system of equations, we can find the demand for the child's leisure time:

$$L^{*} = \frac{Y}{w(1 + \alpha/\beta)} \tag{7}$$

Thus, parents choose the child's amount of leisure time according to the demand function above. This leisure time may be spent in a variety of ways. The child could be engaging in play, whether that be functional/practice play, symbolic play, games-with-rules, social play, or other types of play. They could also spend their leisure time studying for school or otherwise preparing for their future. Each of these different uses of leisure time could have different implications for educational outcomes.

Educational outcomes can be modeled as the output of a standard production function. Glewwe and Muralidharan (2016) model the production function for educational outcomes as

$$A = a(S, Q, C, H, I) \tag{8}$$

where A is skills learned (achievement), S is years of schooling, Q is a vector of school and teacher characteristics (quality), C is a vector of child characteristics (which includes "innate ability"), H is a vector of household characteristics (like credit constraints and parental taste for schooling), and I is a vector of school inputs that are under the control of households (like attendance, effort in school and homework, and school supplies). We can think of study time, which is one of the possible uses of a child's leisure time in the Cobb-Douglass function above, as part of vector I.

Leisure time in the form of play may also impact educational outcomes – as suggested by the psychological literature, play in and of itself may be beneficial to children's cognitive development and educational outcomes. Thus, we can add onto the model of Glewwe and Muralidharan:

$$A = a(S, Q, C, H, I, P) \tag{9}$$

where P is a vector of play characteristics including time spent playing, the type of play, who it was with, etc.

Glewwe and Muralidharan caution that variation in observed school and teacher characteristics (Q) and household characteristics (H) are likely to be correlated with omitted school, teacher, and household variables that determine learning outcomes, leading to biased estimates. In the next section, I discuss my empirical model and the steps I took to reduce this bias.

3 Empirical Model

Drawing on the theoretical model above, we can empirically model the relationship between educational attainment and time spent playing as

$$A_i = \alpha + \beta(T_i) + \delta(X_i) + \epsilon_i$$

where A_i is educational outcomes, T_i is proportion of time spent playing for individual child i at age five, X_i is a vector of controls, and ϵ_i is the error term. I focused on time spent playing as my independent variable of interest because 1) it was a gap in the literature that I could fill, and 2) data availability.

I measured four different educational outcomes (A). Three of these outcomes were test scores: I used the child's test scores in 2016 (age 15) in math, reading, and vocabulary. The fourth outcome was grade attained by 2016.[1]

This model does not demonstrate a neat causal relationship. While it is unlikely that higher test scores from the future caused changes in time spent playing in the past (reverse causality), it is still entirely possible that time spent playing was correlated with the error term due to omitted variables, and thus the regressions produced biased estimates. Omitting variables only results in biased coefficients if the omitted variables are correlated with the included variables – thus, I chose to add control variables that would likely be correlated with time spent playing. Additionally, if the omitted variable is not an important determinant of the dependent variable, the bias from the omitted variable will be small – thus, I only added

[1]Since I controlled for the year the child entered primary school, this outcome essentially measures whether the child progressed steadily through school since starting (whatever year they started), or whether they were held back a grade/dropped out.

control variables that could have a theoretical influence on the child's educational outcomes. By carefully selecting control variables, I was able to mitigate, but not fully eliminate, the omitted variable bias.

One potential source of bias is that children in wealthier households may have been more enriched by play activities. Their play time might have included cognitively stimulating toys, or there may have been things about their home environment that enriched their play experience. Conversely, children in poorer households may not have experienced types of play that improved their cognitive development. In other words, there may have been a qualitative difference between the types of play experienced between wealthier and poorer households, causing an over-estimate of the effect of play time on educational outcomes. In order to address this, I both controlled for wealth in my main regression and performed secondary regressions for each wealth quartile.

Another potential source of bias is that parents with more education may have played with their children more. Since the literature suggests that parent-child play can be extremely beneficial for a child's cognitive development, children with highly-educated parents may have been able to play more with their parent, and thus attain better educational outcomes than children with less-educated parents. However, it is also possible that parents with more education played with their children less due to higher opportunity costs of their time – and thus the children with less-educated parents might actually have reaped more benefits from play. Overall, the actual effect of omitting parental education is ambiguous, but parental education still could easily have impacted the results. In order to address this, I included an interaction term between the share of time spent playing and whether the primary caregiver completed secondary education. This way, we can see what the relationship between play and educational outcomes was specifically for children with highly-educated parents.

A third potential source of bias is that the quality of play may have been different in different environments – rural children's play may have been more (or less) enriching than urban children's play. While the direction of this effect is ambiguous, it could have easily impacted and biased my results. To address this, I included an interaction term between the share of time spent playing and whether the child was living in a rural or urban area at age 5. I also performed secondary regressions for each environment.

In order to align with the claim in psychological literature that different types of play have different types of effects, I included variables that might have influenced the type of play that the child had. I first included whether the child's household owned a television. If a child's household owned a television, the child may have spent some of their play time watching television, which may have been less enriching than other types of play available to the child. Omitting this variable would likely lead to an underestimate of the actual effect of time spent playing on educational outcomes; luckily, I was able to include this as a control variable. I also included whether the child had a sibling aged four through ten when the child was five years old. Omitting this variable would lead to an underestimate of the actual effect of time spent playing on educational outcomes – having a sibling around may have encouraged more social and pretend play, thus improving the quality of the play.

I included several school-related variables to minimize omitted variable bias. Children who started school earlier were likely to have higher test scores and be in a higher grade. Omitting the year the child started school would likely lead to an overestimate of the effect of time spent playing, so this was an important control. The next control I used is whether or not the child was in any school (daycare, preschool, or primary school) when the time use data was taken. If they were in school during this time, one would expect that they played for a much lower share of their time. But for those children that were in school at that time, it is possible that playing more still had an impact on their later test scores or educational attainment – and excluding this variable would lead to an underestimate of the true effect of time spent playing on educational outcomes. Thus, I controlled for whether they were in school when the time use data were collected. I also controlled for the type of school the child attended. Not all schools offered the same quality of education, and educational outcomes were likely different for different types of schools. The schools all fell into one of four categories: public, private, parochial/NGO, and other. Since children often attended several different types of schools throughout their years of education, I measured this with the proportion of school years in each type of school. School type could signal parental investment and interest in the child, so omitting these variables could introduce bias, although the direction of this bias is somewhat ambiguous. I included school type in all my regressions to account for this potential bias.

I also included several demographic variables to minimize bias. Omitting whether the child had an indigenous ethnicity and whether the child's first language was indigenous may likely lead to an underestimate of the true effect of time spent playing, since indigenous children typically have worse educational outcomes in Peru. Omitting the child dependency ratio may lead to an underestimate of the true effect of time spent playing, as more children per adult was likely positively associated with amount of time spent playing, but negatively associated with educational outcomes. Finally, omitting the gender of the child could introduce bias if one gender played more than the other, and also had different academic outcomes due to gender and not due to play. To minimize bias, I included all these demographic variables. I included interaction terms for these demographic variables as well, since they could all have impacted the type of play experienced by the child, and thus the relationship between time spent playing and educational outcomes.

Finally, I included regional dummy variables, and clustered my standard errors by region (more specifically, the sentinel site). This should help account for differences between regions not already captured by the other variables.

Unfortunately, there are a few variables that I was not able to include. One important one is the child's personality. It is quite possible that children with certain personalities played more than other children, and that these children had better (or worse) educational outcomes due to this personality, and not due to the time they spent playing. I would expect this to bias my estimates upwards, but due to the scant literature on play time, it is difficult to say.

I also was not able to include the parents' level of investment and interest in the child's education. Some parents value education more than others, and this value placed on education could easily have led to both less time playing and higher educational outcomes for reasons unrelated to play, and thus my estimate of the importance of time spent playing would be an underestimate. Of course, it is also possible that the parents that value education also gave their child more time to play, which would then lead to an overestimate of the importance of time spent playing; thus, parental value of education could have a somewhat ambiguous effect on the direction of the bias.

Finally, I also was not able to include information about how the child spent their play time. This is particularly frustrating because many studies have shown the importance of parents playing with their children, but I did not have the information to identify whether this was occurring. While this is not strictly necessary to answer the question of how playing more in general was related to educational outcomes, omitting this information may bias my estimates downwards, since I included all types of play, not just those that were beneficial for educational outcomes.

The direction of the overall bias from omitted variables on my estimate of time spent playing is somewhat ambiguous. I would predict that omitting information about the child's personality would lead to an overestimate of the impact of time spent playing, but this relationship is not well-studied, as I noted earlier. I would also predict that omitting information about the parents' preferences for their child's education would lead to an underestimate of the impact of time spent playing, but it is also possible for this relationship to work in the other direction. Overall, there is not a clear direction of the bias of my estimates one way or another. Caution should be used when interpreting my results, due to these important omitted variables.

4 Data and Descriptive Statistics

The Young Lives Survey[2] is a large-scale longitudinal study following children and youth around the world from 2002 to 2016 (Morrow, 2017).[3] In Peru, there have been five rounds of surveys following about 2000 children. The children were selected using sentinel sites, which is a form of purposive sampling from health surveillance studies where the site represents a certain type of population and is thus expected to show trends affecting those people

[2]Data: Boyden (2018a,b); Jones and Huttly (2018); Sanchez et al. (2018); Woldehanna et al. (2018)

[3]The Young Lives Survey has robust ethical practices, with informed consent and reciprocity. For more about Young Lives Survey's ethics, see Morrow (2013).

or areas (Young Lives, 2018). To choose these twenty sites, researchers removed the top five percent in an aggregated measure of access to services, schooling, infant mortality, etc. of the 1,818 districts in Peru at the time, then selected districts semi-randomly to cover urban, peri-urban, rural, coastal, mountain and Amazon areas. Once a district was chosen, a random population center (like a village or hamlet) was chosen within a district, then a random census tract within the population center, then a random street block within the census tract. Fieldworkers visited all dwellings in each block or cluster of houses to find children of the right ages. Once one block was completed, the fieldworker visited a neighboring block, and continued this process until the required number of children (approximately 100 for the younger cohort) was reached. While this process does not necessarily yield a nationally representative sample, the demographic characteristics of their sample mirror those of the DHS, once one takes into account that the probability of a district being selected was proportional to its population size.

Basic summary statistics for my variables of interest and my control variables are shown in Tables 1 and 2. I will next delve into the specifics of my most important variables: time use and educational outcomes.

4.1 Time Use

When the child was around five years old, the Young Lives Survey asked the primary caregiver of the child how much time the child spent on a pre-specified set of activities. The exact question was, "In the last week on a typical work (or school) day how many hours did the child spend on the following activities?" The categories available were sleeping, looking after others (younger siblings, sick people, other household members), domestic tasks (fetching water, firewood, cleaning, cooking, washing, shopping, etc.), unpaid work on the family farm/cattle herding/shepherding/other family business, paid (remunerated) work or activities outside of the household or for someone not in the household, at school, studying outside of school time, and play time/leisure time. Unfortunately, respondents were not able to add their own categories or add nuance to their answers, so it is possible that the child spent time doing other (unreported) activities, or that the child was doing multiple activities at once (like studying and looking after others). For these reasons, along with possible data errors, not all of the time use categories added up to exactly 24 for each child. To adjust for the different total reported hours, I used the share of hours the child spent on each activity (of the total hours reported for that child, what percent of the time did they spend doing x activity?). Table 1 includes basic information about the share of time the child spent on each activity.

Because the survey put play and leisure in the same category, I will use the terms interchangeably from here on out, but it is important to note that they are not necessarily the same – while play is leisure, leisure is not necessarily play. For example, the child may have spent their leisure time watching television, reading for fun, or even simply watching other children play – and these activities are not typically described as play. I tried to account for this somewhat by including the presence of a television in the household as an interaction term in my regressions, but I was not able to fully differentiate between the time a child spent on non-play leisure activities and the time spent on play.

Density plots of time use when the child is five years old (survey round two) are displayed in Figure 1. We can see that play was fairly normally distributed, even as other time use categories had more uneven distributions. In particular, school had two "humps" – one for children who were in school during the time use survey, and one for children who were not.

249 children (about 17%) in my sample were not in school during the time use survey (see Table 2) – this underscores the importance of using whether the child was in school during the time use survey as a control.

What was the true opportunity cost of play? The two time use categories most strongly correlated with play were school and studying (see Table 7 for details). This aligns with the prior theory we laid out – that when a child is not playing, they are spending their time doing school-related activities. This tradeoff is visualized in Figure 8 as well – children who played the most spent less time on all other categories, but most notably on school and studying.

4.2 Educational Outcomes

In 2016 (age 15), the survey included tests for the youth in mathematics, reading, and vocabulary (with the Peabody Picture Vocabulary Test, or PPVT)[4], and these are three of the main outcomes I examine.[5] The fourth measure, the child's grade in 2016, is a measure of educational persistence – whether the child is progressing through school as expected, both by being in school at all and by being in the right grade. Summary statistics for these outcomes are shown in Table 1.

Figure 2 shows density plots of cognitive scores in math, reading and vocabulary, as well as the child's grade in 2016. If the child started primary school in 2008 (around age six or

[4]The Spanish version of the Peabody Picture Vocabulary Test (PPVT) was used for the vocabulary test.

[5]All children were given the same tests. The z-score for these tests was calculated using the full sample of children who took the tests in round five; still, the summary statistics show means only slightly below zero and standard deviations of very slightly below 1, so it is reasonable to continue to interpret these as z-scores.

seven) and did not repeat or skip any grades, they would be expected to be in Secondary Grade 3 ("9" in the plots below) or above in 2016. The school year in Peru runs from around early March to late November, so the child was most likely in the same grade throughout the entire year. There was generally a normal distribution for all three test scores and for expected grade, although vocabulary seemed to have two "peaks", perhaps relating to how some children started school later and so had not learned as much vocabulary.

Figure 3 and Figure 4 show the same educational outcomes, but now examine the relationship with time spent playing. Figure 3 and Figure 4 differ in the sample used to create the density plots. Because not all children were in school yet when their time use data was taken (and the children in school spent less time spent playing since they now spent time at school), it did not make sense to analyze measures of play quantity for both of these groups together. I controlled for whether or not the child was in school by creating density plots for both groups.

I separated out the children who played the most (Upper 30%) and children who played the least (Lower 30%). In both figures we can see that children who played more generally had higher test scores, but the relation between time spent playing and the child's grade in 2016 is somewhat ambiguous, and possibly even negative for children who were in school when the time use data was taken.

Figure 5 and Figure 6 show these same educational outcomes but examine the relationship with wealth. Wealth was measured by the wealth index created by the Young Lives Survey, which is a score between zero and one and is evenly weighted between three categories: housing quality, access to services, and consumer durables (Briones, 2017). The housing quality category itself was evenly weighted between four subcategories: main material of walls, main material of roof, main material of floor, and household density.[6] The access to services category was also evenly weighted between four subcategories: electricity, drinking water source, sanitation facility, and fuel for cooking.[7] The consumer durables category was the total number of certain important items the family possesses.[8]

Figure 5 and Figure 6 separate out the children from the most and least wealthy 30% of households and examine how their educational outcomes differed. Again, I controlled for whether the child was in school by creating these density plots by wealth for both groups (children in school, and children not in school). It is clear from both figures that children in wealthier families tended to have higher test scores and be in a higher grade in 2016.

Figure 7 illustrates the relationship between time spent playing and wealth. I separated out the children from the most and least wealthy 30% of households, then examined how the amount of time they spent playing differed. We can see that wealthier children spent more

[6]The main material of walls received a score of 1 if it was made of brick/concrete or of concrete blocks, and a score of 0 for anything else. The main material of roof received a score of 1 if it was made of concrete/cement, galvanized/corrugated iron, or tiles/slates, and a score of 0 for anything else. The main material of the floor received a score of 1 if it was made of cement/tile, laminated material, stone (granite/marble), polished stone, or parquet, and a score of 0 for anything else. Household density was calculated as the ratio of the number of rooms (excluding kitchen, bathrooms, corridor, and garage) and the number of household members; this number was rescaled to be between 0 and 1, compared with the maximum and minimum household densities in the sample for that country.

[7]Electricity received a 1 if they had electricity (not counting car batteries). Drinking water received a 1 if it was piped water to the house/plot (public network) or a well/tube well with a hand pump. Sanitation received a 1 if they had a flush toilet/septic tank or a pit latrine. Fuel for cooking received a 1 if it was gas/electricity or kerosene/paraffin.

[8]There were 12 consumer durables items: radio, television, bicycle, motorbike, automobile, landline phone, mobile phone, refrigerator, stove, blender, iron, and record player.

time playing than poorer children, but that the distributions are not extremely different. Caution should be used when interpreting the impact of time spent playing on educational outcomes, since there was no such thing as a wealthy child who spent none of their time playing.

5 Results

I performed several basic linear regressions to understand the relationship between play and educational outcomes. The independent variable of interest in all these regressions is the share of time spent on play. The dependent variables are math z-scores, reading z-scores, vocabulary z-scores, and the student's grade in 2016.

5.1 Share of time spent playing

Table 3 reports the results from the first four regressions, where the independent variable of interest is the share of time spent playing, and a separate ordinary least squares (OLS) regression was performed for each of the four of the educational outcome variables, using all the controls and interactions specified earlier. For detailed results, see Table 8. I did not observe a statistically significant relationship between play at age five and educational outcomes in 2016 (age fifteen). However, some of the interaction terms were statistically significant, and I will briefly mention them here.

Being male had a strong positive relationship with math scores and a strong negative relationship with reading scores. The interaction between share of time spent on play and male was strongly positive for reading scores, however – playing more was associated with higher reading scores for males. However, for a male to have a higher score than a female who spent none of her time playing, he had to spend at least 22.5% of his time playing (slightly over the mean of 20.6%).

More children per adult in a household (higher child dependency ratio) had a strong positive relationship with math scores. This is counterintuitive; controlling for all other household characteristics, one would expect that households with more children would have fewer resources to dedicate to each child, thus lowering their expected outcomes. But the interaction between the share of time spent on play and the child dependency ratio was negative: playing more was related to lower test scores given a child dependency ratio, and having a higher child dependency ratio was related to lower test scores given a set share of time spent playing. The ultimate effect of the child dependency ratio is ambiguous and depends on the actual share of time spent on play (the more time spent on play, the more negative the effect of the higher child dependency ratio), and is tricky to interpret due to the continuous nature of both the child dependency ratio and the share of time spent on play.

The interaction with having a similarly-aged sibling, as well as with the household having a working television, was negative and statistically significant for vocabulary scores – I discuss this more at length later.

5.2 Wealth Quartiles

I performed the same analysis, using the same control variables and interaction terms, but separated the children into different groups based on their wealth quartiles. It is possible that play had very different effects depending on where the child's family was on the socio-economic ladder, or that children in different wealth quartiles played differently, with different effects on educational outcomes as a result.

The results of this analysis can be seen in Table 4. For detailed results, see Tables 9-12. There was a strong positive relationship between the share of time spent playing and math scores for children in the lowest wealth quartile. Here, spending 10% more time on play (increasing the share of time spent on play by 0.1) at age five was associated with a 0.239 increase in the child's math z-score. The mean math z-score for children in the lowest wealth quartile was -.396, with a standard deviation of about .042. For this group of children, then, a 0.239 increase in the z-score was substantial, but not enough to bring them up to the average of all the other children sampled.

The reason for the play-math relationship for children in the lowest wealth quartile is unclear. Only 18% of households in the lowest wealth quartile had televisions (compared to 69%, 92%, and 99% of quartiles two, three, and four respectively), so it is possible that those children were spending play time not watching television, and instead engaging in more developmentally useful play than children in other wealth quartiles. But playing more did not seem to have a strong relationship with any other educational outcomes for children in the lowest wealth quartile, so while this is a strong, statistically significant result, caution must be used in drawing conclusions from it, barring other corroborating work.

There was also a strong negative relationship between the share of time spent playing and educational attainment in 2016 for children in the second-lowest wealth quartile, independent of the year they began primary school. Spending 10% more time on play at age five was associated with a 0.3113 decrease in the child's educational attainment in 2016. The mean educational attainment for children in the second wealth quartile was 8.804 (so between eight and nine years of schooling completed), with a standard deviation of 0.056. While a

0.3 decrease is not practically meaningful, the negative relationship may still be something to consider.

The reason for the play-vocabulary relationship for children in the second wealth quartile is also unclear. It is possible that something in their play environment made them less likely to engage in the types of play that have strong effects on child development. However, I could not discern what aspects of their environment that might be.

5.3 Rural and Urban Environments

I again performed the same analysis, using the same control variables and interaction terms, but separated the children into different groups based on whether they were in an urban or rural setting at age 5. Again, it is possible that play had very different effects depending on the child's locational context, or that children in urban and rural areas played differently, with different effects on educational outcomes as a result.

Ultimately, I did not see any evidence of this - the results of the rural/urban analysis can be seen in Table 5. For detailed results, see Tables 13 and 14. There was no statistically significant relationship between play at age five and any of the educational outcomes, for either the rural or urban children. This is not necessarily evidence that children's experiences of play were not different in urban and rural areas, but it is evidence that any differences in their types of play did not impact educational outcomes.

5.4 Types of Play

As discussed earlier, the psychological literature examines many different types of play. I did not have much information about the type of play the children were engaging in during their time spent playing, but I did know some things about the household that might have influenced what the child was doing with their leisure time. I knew if a child's household owned a television – and a child spending their time watching television may have been less enriching than the child spending their time engaging in pretend play, for example. In the main regression that included all the children (Table 3), owning a television had a strong positive relationship with vocabulary scores, but the interaction with the share of time spent on play was negative, which implies that the child actually watching TV had negative effects on their vocabulary scores. This would then imply that some types of play are neutral towards test scores, while some are actually harmful for test scores.

I also had information about whether the child had siblings in a relatively similar age group (age four to age ten, when the main child was age five). It is possible that having a sibling around encouraged more social and pretend play. In my main regression (Table 3), having a sibling with a similar age had a negative relationship with math scores. Meanwhile, the interaction between share of time spent on play and having a sibling was negative for vocabulary scores. So siblings were negatively related to math test scores (independent of the child dependency ratio!), and spending more time playing and having a sibling had a negative relationship with vocabulary test scores. Both of these results imply that siblings, and spending time playing with siblings, may actually have a detrimental effect on educational outcomes. Perhaps social play is not as useful as previously hypothesized, but it could also be that social play is critical for child development in ways unrelated to educational outcomes. And it should also be noted that having a sibling does not necessarily result in social play – I still did not have information about what the child was doing during the time that their caregivers reported as play.

5.5 Important Covariates

Several of the variables I controlled for were extremely important determinants of educational outcomes. The most notable was household wealth, which had a positive and statistically significant relationship with all the outcomes I measured in my main regression. It also had a positive and statistically significant impact on all educational outcomes for rural children, and for the math and vocabulary scores of urban children. This reflects what we saw earlier in Figures 5 and 6 – wealthier children have better educational outcomes than poorer children. This result is intuitive, as wealthier households can invest more in their children than poorer households can, and thus their children have better outcomes. This overwhelming result is evidence that education is not a "great equalizer" in Peru – students from wealthier families have better educational outcomes, and policymakers should pay special attention to income inequality in schools in order to provide as many opportunities as possible to students from poor families.

Several variables related to schools also had a strong relationship with educational outcomes. Starting school in a later year had a negative relationship with all test scores. This is unsurprising, since one learns more content in math, reading, and vocabulary as one continues in school, and thus would likely score higher in those subject areas the longer one was in school. Relatedly, spending more years in a private school had a strong positive relationship with all test scores. Perhaps this is because the quality of education at these schools was generally higher, but it could also be that there are qualities correlated with a child going to private school that are the main factor in test scores (like high-ability or wealthy children being sent to private school).

Several variables related to family characteristics also had a strong relationship with educational outcomes. The level of education of the primary caregiver had a strong positive relationship with math and vocabulary scores, and a positive but not statistically significant relationship with reading scores and grade in 2016. This positive relationship aligns with intuition and prior studies – children of educated parents go farther and do better in education, since educated parents often value education more than their less-educated counterparts. If the child spoke an indigenous language as a first language, they performed statistically significantly worse in the vocabulary tests than children with Spanish as a first language. This makes sense given that the vocabulary test was given in Spanish.

5.6 Attrition

Some children chosen for the first round of the survey moved away, refused to continue participation, or were otherwise not surveyed in subsequent rounds – the attrition rate between rounds one and five was 8.2%. In order to understand how random or non-random this attrition was, I examined the determinants of attrition. I performed three probit regressions with three dependent variables: whether the child was ever missing from a survey round, whether the child was missing in survey round two (when the time use data was taken), and whether the child was missing in survey round five (when the educational outcomes were measured).

The results from these regressions can be seen in Table 6. The only factor that ever statistically significantly predicted attrition was whether the primary caregiver felt "part of the community" in the first round. The direction of this relationship was intuitive – people who felt like they were a part of the community were less likely to attrit. But beyond that, I saw little evidence of a systematic trend through which children dropped out of the study, and so I do not consider attrition to be a major limitation in my analysis.

5.7 Limitations

While all the primary caregivers were asked the same question about how many hours the child spent on play, they may not have all interpreted the question in the same way. The definition of play is somewhat nebulous, so different caregivers might have categorized the same actions (e.g. watching TV) differently. This could definitely impact my results, especially if similar households categorized play similarly. Thus, because of the subjective nature of play, the time use survey data that I use may be unreliable and render unreliable results.

Social desirability bias may have also impacted responses to the time use questions. If a caregiver saw it socially desirable to say, for example, that their child was attending school even when they were not, the caregiver may have altered their responses to match what they perceived as socially desirable. This would also render the time use survey data unreliable and bring the accuracy of my results into question.

While I included numerous controls, omitted variable bias still may have been present. If that was the case, then my estimates of the relationship between time spent playing and educational outcomes may be inaccurate. One important variable I did not control for is the child's personality – children with certain personalities related to higher (or lower) educational outcomes might have selected into (or out of) play when they were children. I think it reasonable to say that the child's play time is mostly determined by the primary caregiver, but this was likely still a factor. I also was not able to control for the parents' level of interest and investment in their child's education – parents who highly valued education may have allowed less time for their child to play, but also provided many other inputs important to the education production function. While I did control for the type of school the child attended and the primary caregiver's level of education (which may be proxies for how much the parent cared about the child's education), I was not able to fully address this problem.

Play is a complicated concept, and the time a child spent on "play time/leisure time" may not truly get at the positive aspects of play that researchers have previously identified. For example, pretending to be a shopkeeper, playing hopscotch, and watching television would all have been counted in the same category of play in this survey. But many would categorize these as different types of play, with different types of potential effects on educational outcomes. While I do include the presence of a television in the home and whether there were siblings in a similar age range in my regressions as controls and as interactions with play, there are other kinds of play experiences the children may have had that I did not address. In particular, many studies have shown the importance of parents playing with the child, but I did not have the information to identify whether or not this was occurring.

6 Conclusion

Understanding the effects of time spent playing on children's educational outcomes would allow us to make the best recommendations for how parents should allocate their children's time. It would also settle a theoretical debate between economists and psychologists that hinges on whether play has inherent value. I found little evidence that spending more time playing at age five was related to educational outcomes at age fifteen. However, I did find that playing more was related to an increased math z-score for children in the lowest wealth quartile, and lower educational attainment for children in the second-lowest wealth quartile. I also found that playing more had a strong positive relationship with reading scores for male children, and that playing more had a strong negative relationship with math scores as the child dependency ratio increased. Finally, I have some evidence that spending leisure time watching television or playing with siblings may have actually negatively impacted educational outcomes. In general, these findings suggest that there may be a relationship between play and educational outcomes, but perhaps only for particular populations. Further study is needed to carefully untangle these relationships and settle this debate.

Even as the relation between play quantity and educational outcomes remains hazy, this work solidifies the importance of household wealth in determining a child's educational outcomes. In nearly all of my regressions, wealth had a positive and statistically significant impact on educational outcomes. As we push for policies that improve children's educational outcomes and lives, we must be cognizant of the role wealth plays in determining these outcomes, and actively work to provide opportunities to children from poor families in order to narrow this educational divide.

References

Abimpaye, M., Dusabe, C., Nzabonimpa, J. P., Ashford, R., and Pisani, L. (2019). Improving parenting practices and development for young children in Rwanda: Results from a randomized control trial. *International Journal of Behavioral Development*, 44(3). https://doi.org/10.1177/0165025419861173.

Andrietti, V. and Velasco, C. (2015). Lecture attendance, study time, and academic performance: A panel data study. *Journal of Economic Education*, 46(3):239–259. https://doi.org/10.1080/00220485.2015.1040182.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal*, 75(299):493–517. https://doi.org/10.2307/2228949.

Bergen, D. (2018). Infant sensorimotor play: Development of sociocultural competence and enactive cognition. In Smith, P. K. and Roopnarine, J. L. E., editors, *The Cambridge Handbook of Play: Developmental and Disciplinary Perspectives*, pages 125–141. Cambridge University Press. https://doi.org/10.1017/9781108131384.008.

Boyden, J. (2018a). Young lives: an international study of childhood poverty: Round 2, 2006. [data collection]. 3rd Edition. UK Data Service. SN: 6852, http://doi.org/10.5255/UKDA-SN-6852-3.

Boyden, J. (2018b). Young lives: an international study of childhood poverty: Round 3, 2009. [data collection]. 3rd Edition. UK Data Service. SN: 6853, http://doi.org/10.5255/UKDA-SN-6853-3.

Briones, K. (2017). "How many rooms are there in your house?": Constructing the young lives wealth index. *Young Lives*. https://www.younglives.org.uk/content/how-many-rooms-are-there-your-house-constructing-young-lives-wealth-index.

Clark, N. (2015). Education in Peru. Retrieved October 30, 2019 from https://wenr.wes.org/2015/04/education-in-peru.

Crispin, L. M. and Kofoed, M. (2019). Does time to work limit time to play?: Estimating a time allocation model for high school students by household socioeconomic status. *Con-*

temporary Economic Policy, Western Economic Association International, 37(3):524–544. https://doi.org/10.1111/coep.12411.

Cueto, S., Guerrero, G., León, J., Seguin, E., and Muñoz, I. (2009). Explaining and overcoming marginalization in education: A focus on ethnic/language minorities in Peru. In *Paper commissioned for the EFA Global Monitoring Report 2010*. Reaching the marginalized. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000186589.

Finnish National Agency for Education (2018). Compulsory education in Finland. https://www.oph.fi/en/statistics-and-publications/publications/compulsory-education-finland.

Garces, E., Thomas, D., and Currie, J. (2002). Longer-term effects of Head Start. *American Economic Review*, 92(4):999–1012. https://doi.org/10.1257/00028280260344560.

Gershuny, J. (2009). Veblen in reverse: Evidence from the Multinational Time-Use Archive. *Social Indicators Research*, 93(1):37–45. https://doi.org/10.1007/s11205-008-9369-6.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M., and Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, 344(6187):998–1001. https://doi.org/10.1126/science.1251178.

Glewwe, P., Jacoby, H. G., and King, E. M. (2001). Early childhood nutrition and academic achievement: A longitudinal analysis. *Journal of Public Economics*, 81(3):345–368. https://doi.org/10.1016/S0047-2727(00)00118-3.

Glewwe, P. and Muralidharan, K. (2016). Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education (Vol. 5)*, page 653–743. Elsevier. https://doi.org/10.1016/B978-0-444-63459-7.00010-5.

Gould, E. D., Lavy, V., and Paserman, M. D. (2011). Sixty years after the magic carpet ride: The long-run effect of the early childhood environment on social and economic outcomes. *The Review of Economic Studies*, 78(3):938–973. https://doi.org/10.1093/restud/rdq038.

Hancock, L. (2011). Why are Finland's schools successful? *Smithsonian Magazine*. https://www.smithsonianmag.com/innovation/why-are-finlands-schools-successful-49859555/.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1-2):114–128. https://doi.org/10.1016/j.jpubeco.2009.11.001.

Jones, N. and Huttly, S. (2018). Young lives: an international study of childhood poverty: Round 1. [data collection]. 6th Edition. UK Data Service. SN: 5307, http://doi.org/10.5255/UKDA-SN-5307-3.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., and Shonkoff, J. P. n. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences*, 103(27):10155–10162.

Lillard, A. S. (2013). Playful learning and Montessori education. *NAMTA Journal*, 38(2):137–174.

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., and Palmquist, C. M. (2013). The impact of pretend play on children's development: A review of the evidence. *Psychological Bulletin*, 139(1):1–34. https://doi.org/10.1037/a0029321.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R., and Sanbonmatsu, L. (2013). Long-term neighborhood effects on low-income families: Evidence from Moving to Opportunity. NBER Working Paper no. 18772.

Michaels, J. W. and Miethe, T. D. (1989). Academic effort and college grades. *Social Forces*, 68(1):309–319. https://doi.org/10.2307/2579230.

Morrow, V. (2013). Practical ethics in social research with children and families in young lives: A longitudinal study of childhood poverty in Ethiopia, Andhra Pradesh (India), Peru and Vietnam. *Methodological Innovations Online*, 8(2):21–35. https://doi.org/10.4256/mio.2013.011.

Morrow, V. (2017). *A Guide to Young Lives Research.* Young Lives. https://www.younglives.org.uk/content/guide-young-lives-research.

Nicolopoulou, A. (2018). Pretend and social pretend play: Complexities, continuities, and controversies of a research field. In Smith, P. K. and Roopnarine, J. L. E., editors, *The Cambridge Handbook of Play: Developmental and Disciplinary Perspectives*, page 183–199. Cambridge University Press. https://doi.org/10.1017/9781108131384.011.

OECD (2019). Peru country note. Programme for International Student Assessment (PISA). Results from PISA 2018.

Pellegrini, A. D. (2009). *The Role of Play in Human Development.* Oxford University Press.

Pellegrini, A. D., Huberty, P. D., and Jones, I. (1995). The effects of recess timing on children's playground and classroom behaviors. *American Educational Research Journal*, 32(4):845–864. https://doi.org/10.2307/1163338.

Piaget, J. (1962). *Play, dreams, and imitation in childhood*. Norton, New York, NY.

Pike, G. R., Kuh, G. D., and Massa-McKinley, R. C. (2009). First-year students' employment, engagement, and academic achievement: Untangling the relationship between work and grades. *Journal of Student Affairs Research and Practice*, 45(4). https://doi.org/10.2202/1949-6605.2011.

Sanbonmatsu, L., Kling, J. R., Duncan, G. J., and Brooks-Gunn, J. (2006). Neighborhoods and academic achievement results from the Moving to Opportunity experiment. *Journal of Human Resources*, 41(4):649–691. https://doi.org/10.3368/jhr.XLI.4.649.

Sanchez, A., Penny, M., Woldehanna, T., Galab, S., Boyden, J., and Duc, L. T. (2018). Young lives: an international study of childhood poverty: Round 5, 2016. [data collection]. UK Data Service. SN: 8357, http://doi.org/10.5255/UKDA-SN-8357-1.

Schuman, H., Walsh, E., Olson, C., and Etheridge, B. (1985). Effort and reward: The assumption that college grades are affected by quantity of study. *Social Forces*, 63(4):945–966. https://doi.org/10.2307/2578600.

Schweinhart, L. J. and Weikart, D. P. (1997). The High/Scope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly*, 12(2):117–143. https://doi.org/10.1016/S0885-2006(97)90009-0.

Sevilla, A., Gimenez-Nadal, J. I., and Gershuny, J. (2012). Leisure inequality in the United States: 1965—2003. *Demography*, 49(3):939–964.

Shonkoff, J. P., Phillips, D. A., and (Eds.), N. R. C. U. S. (2000). *From neurons to neighborhoods: The science of early child development*. National Academy Press, Washington, D.C.

Tamis-LeMonda, C. S., Shannon, J. D., Cabrera, N. J., and Lamb, M. E. (2004). Fathers and mothers at play with their 2- and 3-year-olds: Contributions to language and cognitive development. *Child Development*, 75(6):1806–1820. https://doi.org/10.1111/j.1467-8624.2004.00818.x.

The Economist (2019). Pisa results can lead policymakers astray. *The Economist*. https://www.economist.com/international/2019/12/05/pisa-results-can-lead-policymakers-astray.

UNESCO Institute for Statistics (2019). UIS statistics. Retrieved October 30, 2019 from http://data.uis.unesco.org/.

Urke, H. B., Contreras, M., and Matanda, D. J. (2018). The influence of maternal and household resources, and parental psychosocial child stimulation on early childhood development: A cross-sectional study of children 36–59 months in Honduras. *International Journal of Environmental Research and Public Health*, 15(5):926. https://doi.org/10.3390/ijerph15050926.

Walker, T. D. (2015). Kindergarten: While American kids read, their Finnish peers play. *The Atlantic*. https://www.theatlantic.com/education/archive/2015/10/the-joyful-illiterate-kindergartners-of-finland/408325/.

Woldehanna, T., Galab, S., Sanchez, A., Penny, M., Duc, L. T., and Boyden, J. (2018). Young lives: an international study of childhood poverty: Round 4, 2013-2014. [data collection]. 2nd Edition. UK Data Service. SN: 7931, http://doi.org/10.5255/UKDA-SN-7931-2.

World Bank (2018a). World development indicators: Educational attainment, at least completed primary, population 25+, total.

World Bank (2018b). World development indicators: Educational attainment, at least completed upper secondary, population 25+, total.

Worlddata.info (2019). Average income around the world. Retrieved October 30, 2019, from https://www.worlddata.info/average-income.php.

Young Lives (2018). Young lives survey design and sampling (round 5): Peru. https://www.younglives.org.uk/node/8583.

Figures

Figure 1: Time Share Density Plots

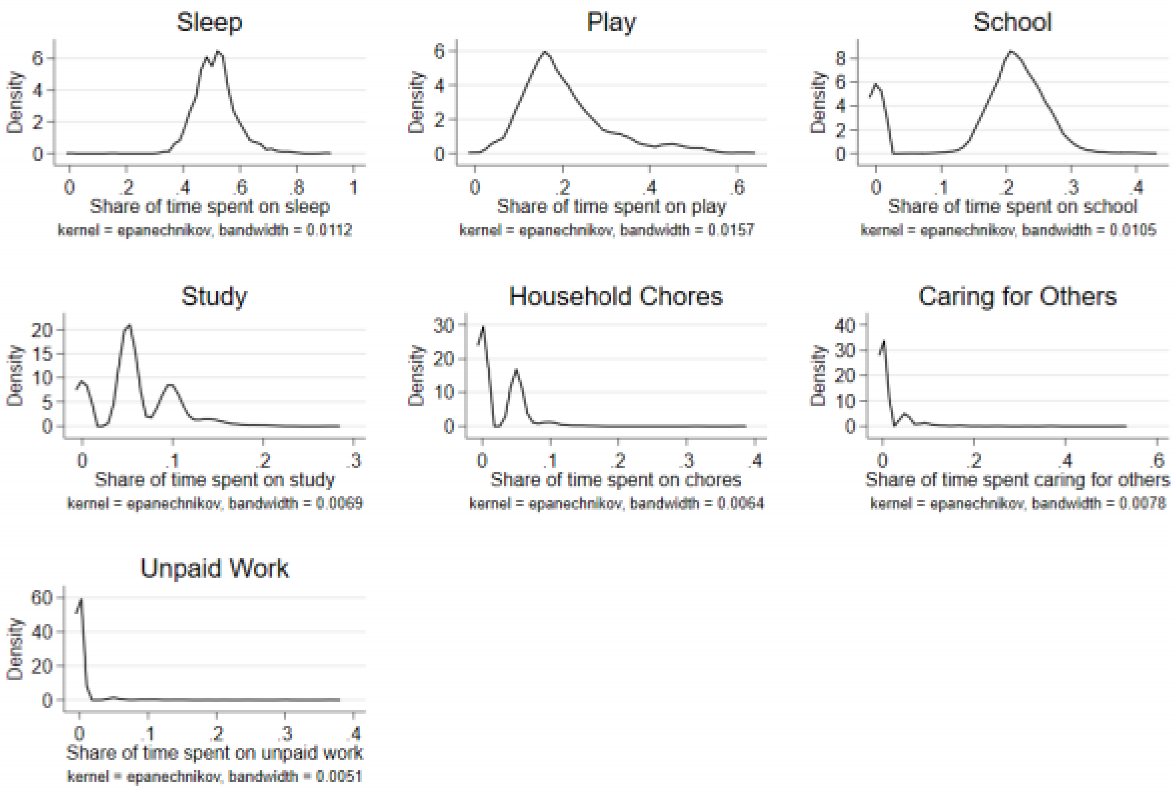


Figure 2: Educational Outcomes Density Plots

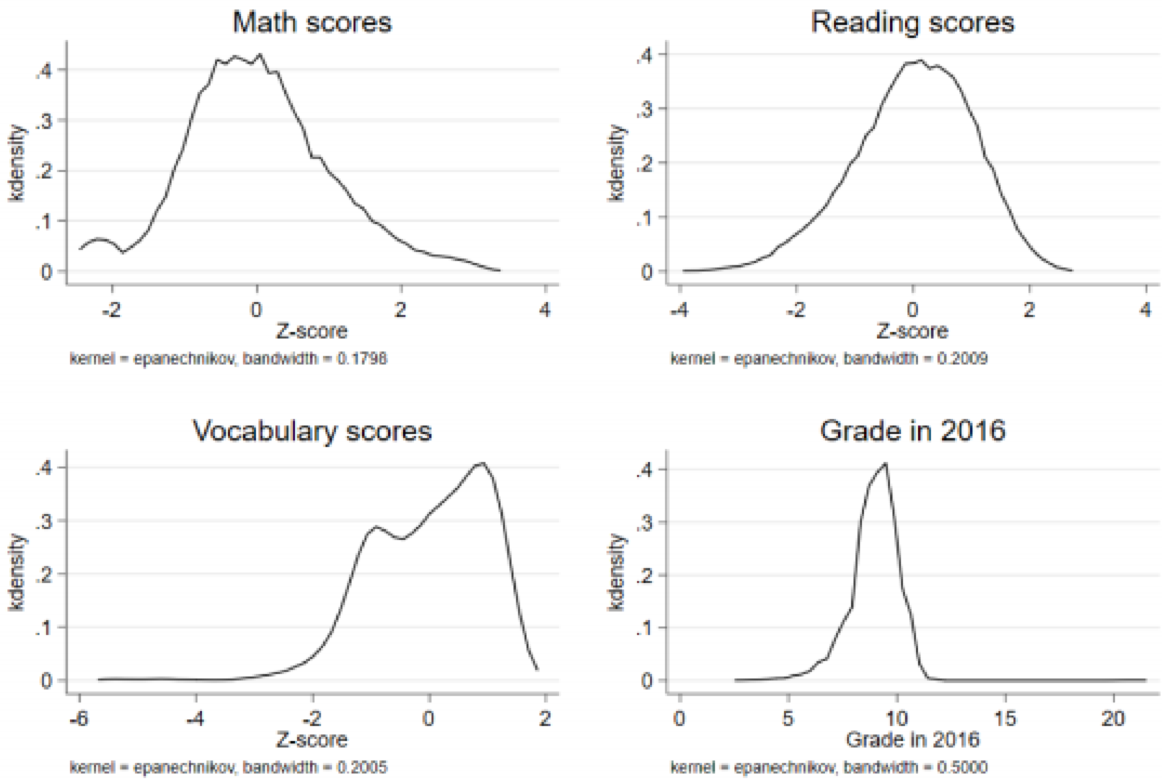


Figure 3: Educational Outcomes for Children in School, by Play

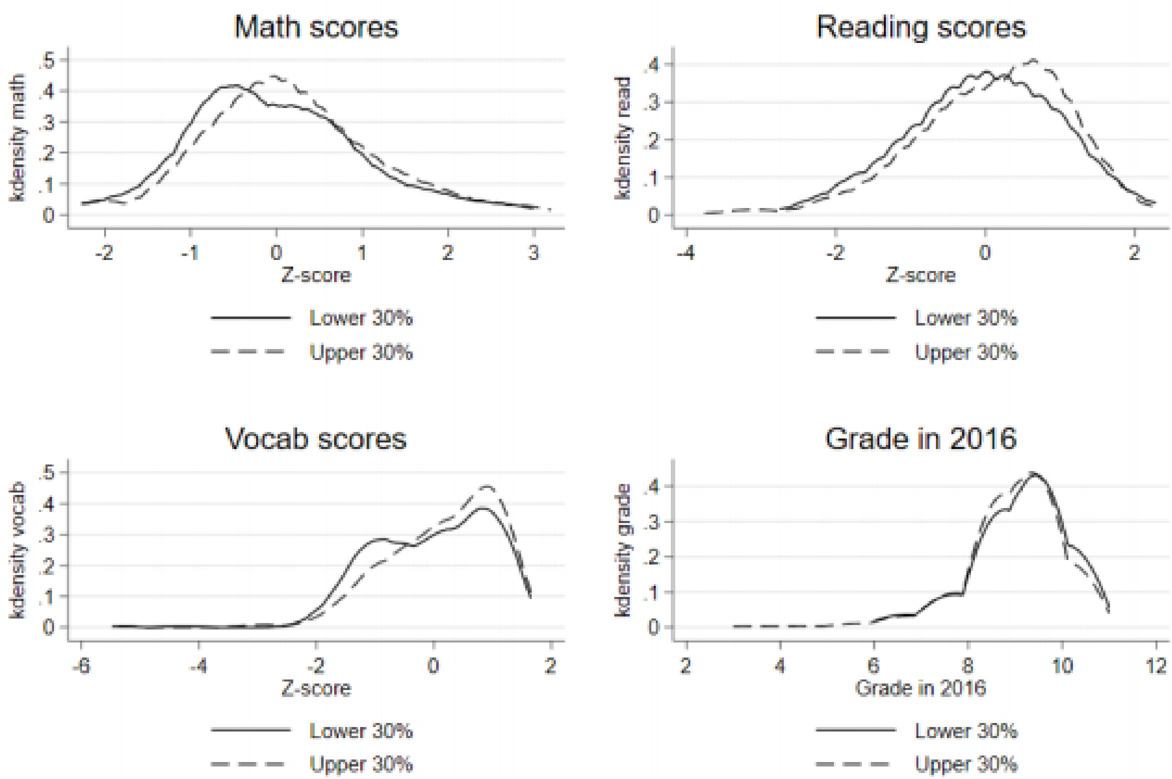


Figure 4: Educational Outcomes for Children Not in School, by Play

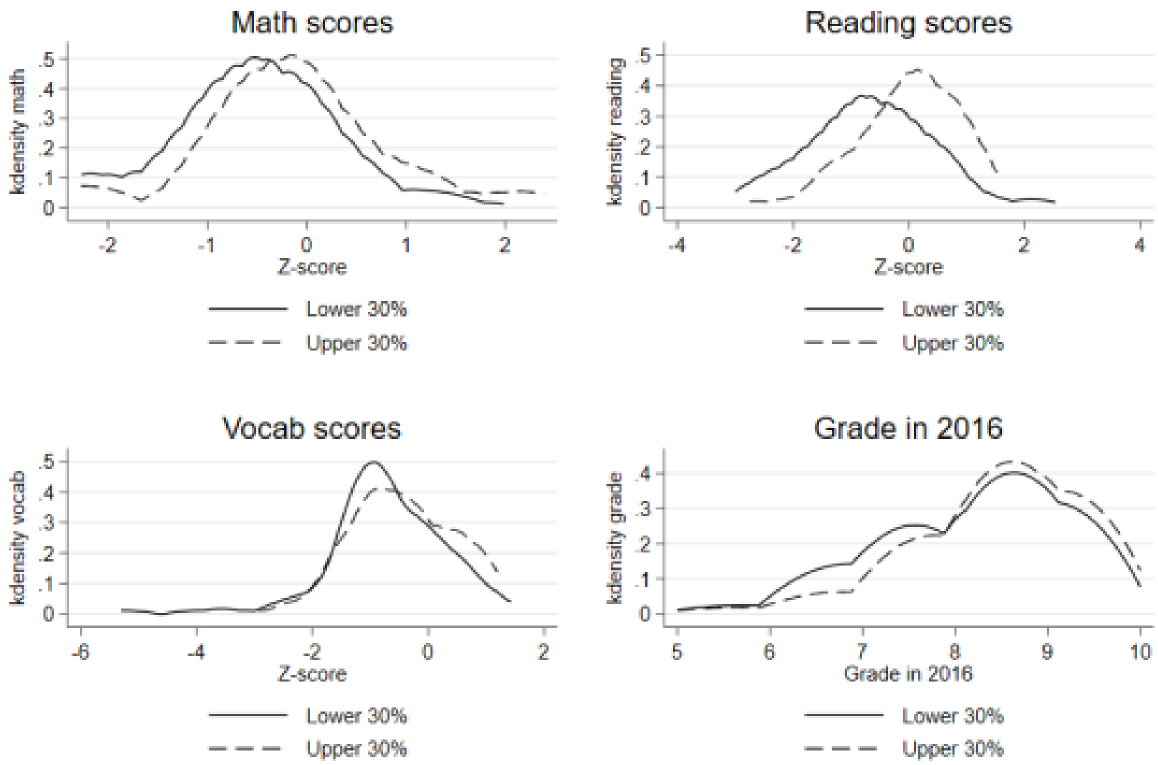


Figure 5: Educational Outcomes for Children in School, by Wealth

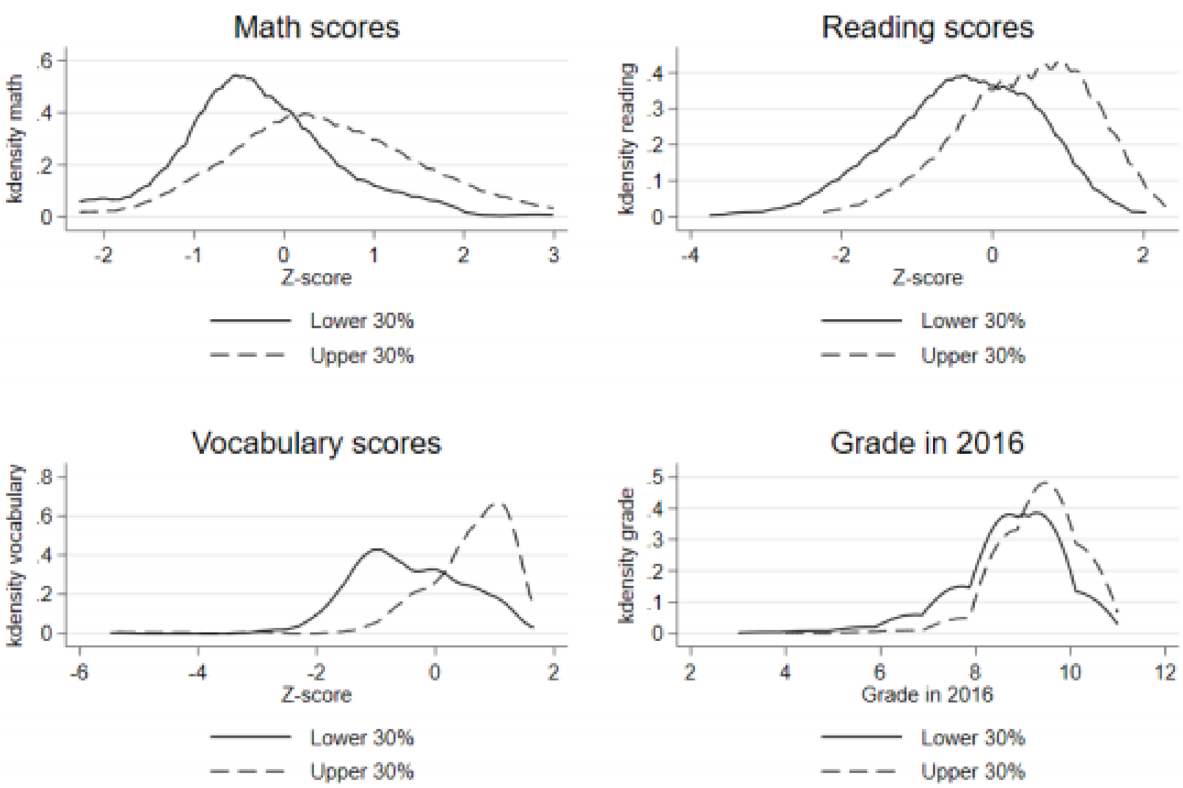
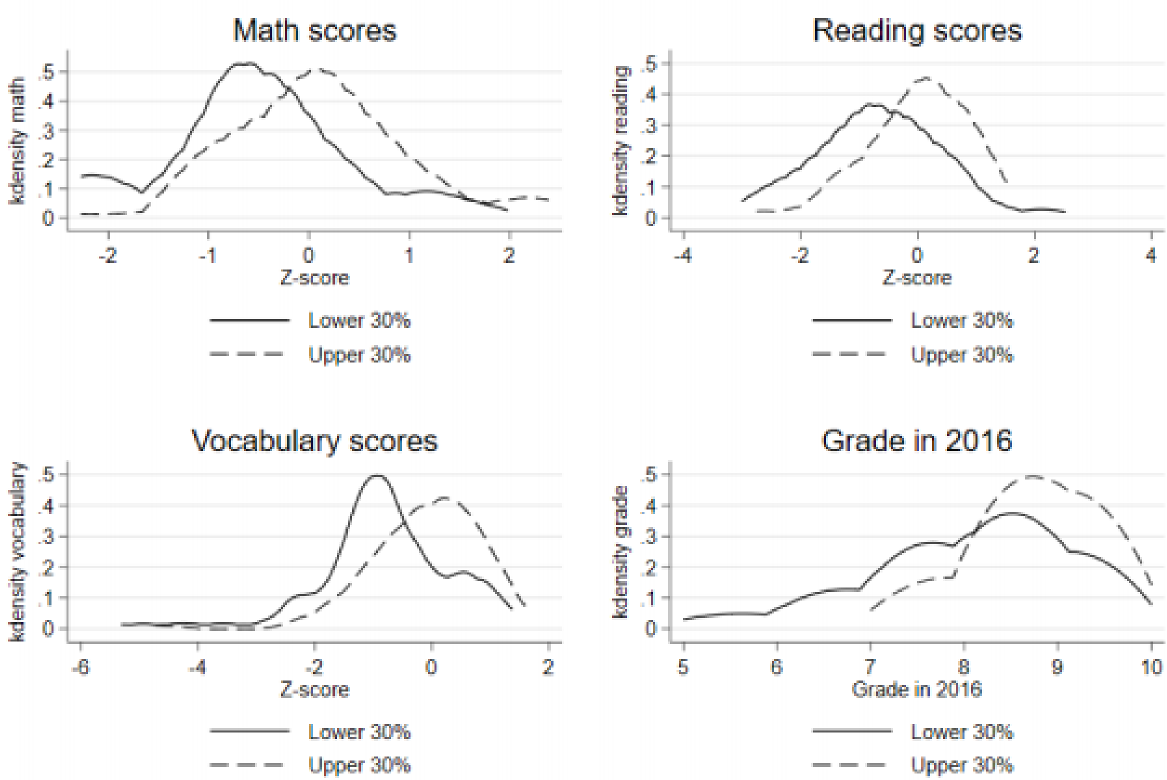


Figure 6: Educational Outcomes for Children Not in School, by Wealth



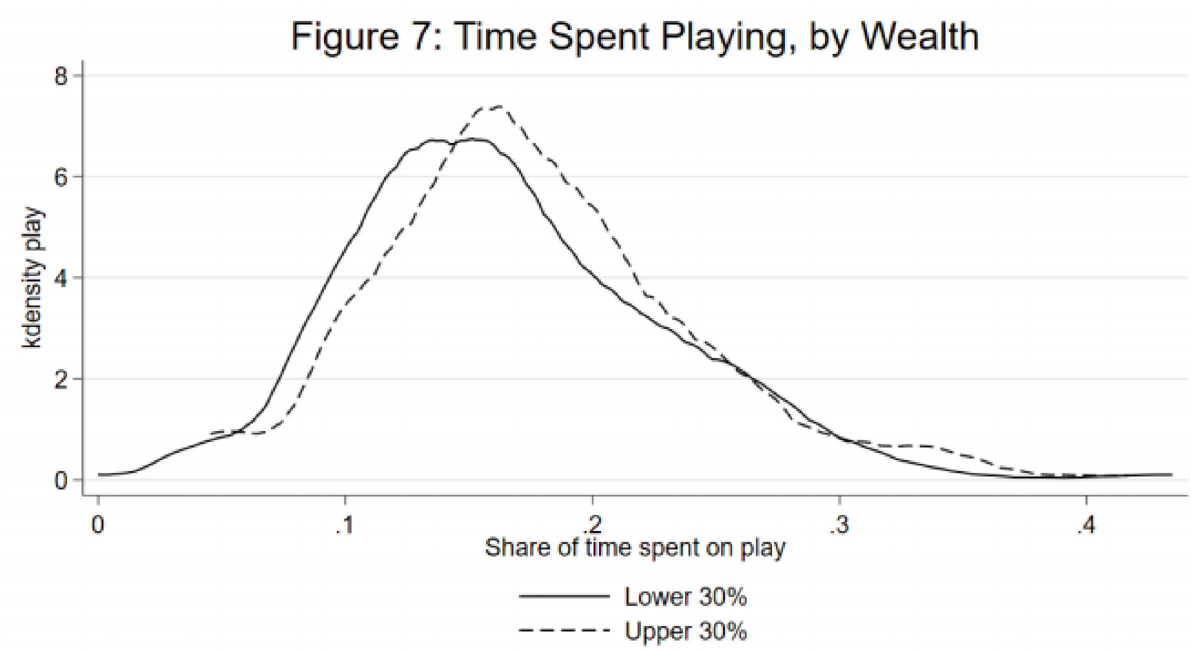


Figure 7: Time Spent Playing, by Wealth
kdensity play
0
2
4
6
8
0
.1
.2
.3
.4
Share of time spent on play
Lower 30%
Upper 30%

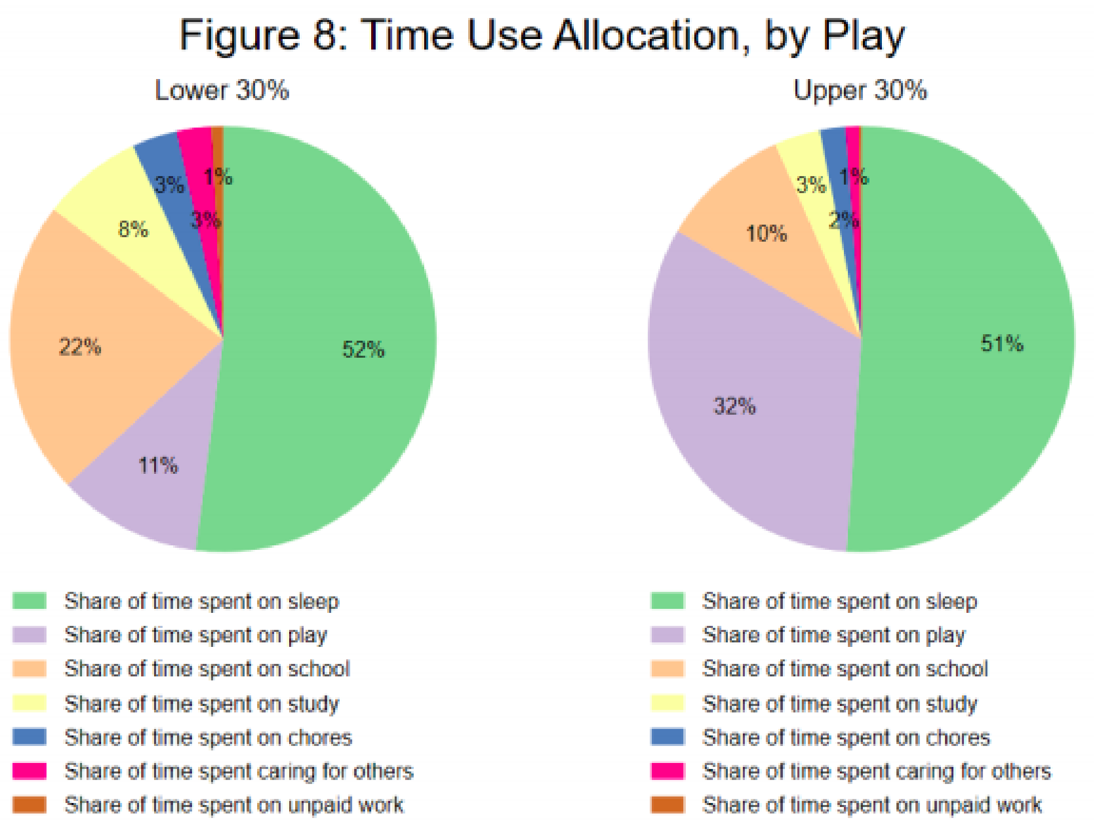


Figure 8: Time Use Allocation, by Play
Lower 30%
52%
11%
22%
8%
3%
3%
1%
Upper 30%
51%
32%
10%
3%
2%
1%
Share of time spent on sleep
Share of time spent on play
Share of time spent on school
Share of time spent on study
Share of time spent on chores
Share of time spent caring for others
Share of time spent on unpaid work

Tables

Table 1: Summary Statistics

VARIABLES	Mean	Std. Dev.	Min	Max
Math z-scores	0.051	0.961	-2.271	3.196
Reading z-scores	0.017	0.990	-3.756	2.535
Vocabulary z-scores	0.025	0.977	-5.486	1.660
Grade in 2016	8.960	1.003	4	21
Share of time spent sleeping	0.512	0.071	0	0.909
Share of time spent playing	0.203	0.100	0	0.625
Share of time spent at school	0.182	0.089	0	0.421
Share of time spent studying	0.060	0.040	0	0.227
Share of time spent on household chores	0.024	0.032	0	0.381
Share of time spent caring for others	0.014	0.038	0	0.526
Share of time spent on unpaid work	0.004	0.027	0	0.375
Share of time spent on paid work	0	0	0	0
Wealth index	0.474	0.229	0.0002	0.922
Year started school	2007.511	0.736	2003	2009
Child dependency ratio	1.228	0.757	0.125	7
Proportion of school years in public schools	0.839	0.285	0	1
Proportion of school years in private schools	0.148	0.278	0	1
Proportion of school years in parochial or NGO schools	0.008	0.045	0	0.571
Proportion of school years in other schools	0.005	0.028	0	0.300

All statistics are for the most restricted sample (1753 observations, limited by data for Grade in 2016).

Table 2: Frequency table

VARIABLES	Yes	No
Male	879	874
Child's first language is indigenous	249	1504
Indigenous ethnicity	272	1481
Rural	533	1220
Had a sibling ages 4-10	832	921
Caregiver completed secondary education	663	1090
Household owns working television	1235	518
In school during time use survey	1459	294

All statistics are for the most restricted sample (1753 observations, limited by data for Grade in 2016).

Table 3: Share of time spent on play and educational outcomes

VARIABLES	(1) Math z-score	(2) Reading z-score	(3) Vocab z-score	(4) Grade in 2016
Share of time spent on play	1.273	0.199	0.874	-1.039
	(0.888)	(0.970)	(0.557)	(0.635)
Share of time spent playing x Child dependency ratio	-0.598***	-0.235	0.112	-0.0129
	(0.191)	(0.401)	(0.241)	(0.240)
Share of time spent playing x Male	0.0855	0.999**	0.584	0.401
	(0.262)	(0.402)	(0.396)	(0.345)
Share of time spent playing x Indigenous first language	-0.00489	1.236	0.561	0.359
	(0.529)	(0.935)	(0.724)	(0.446)
Share of time spent playing x Indigenous ethnicity	0.140	0.214	0.321	-0.0191
	(0.644)	(0.672)	(0.769)	(0.482)
Share of time spent playing x Rural	-0.365	-0.780	-0.709	-0.0600
	(0.514)	(0.562)	(0.649)	(0.298)
Share of time spent playing x Had sibling age 4-10	0.442	0.0227	-0.834**	0.294
	(0.396)	(0.564)	(0.358)	(0.414)
Share of time spent playing x Caregiver completed secondary education	0.413	0.307	-0.241	0.0260
	(0.640)	(0.463)	(0.321)	(0.465)
Share of time spent playing x Household owns working television	-0.328	-0.323	-0.700***	0.371
	(0.459)	(0.490)	(0.190)	(0.328)
Constant	-0.723*	0.358	-0.363	12.61***
	(0.364)	(0.413)	(0.274)	(0.534)
Observations	1,805	1,768	1,780	1,753
R-squared	0.217	0.222	0.354	0.294
Demographic controls	YES	YES	YES	YES
Wealth controls	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note. All estimates were obtained using a linear probability model. Standard errors are clustered by sentinel site. Controls include household wealth, gender, indigenous first language, indigenous ethnicity, primary caregiver education level, owning a working television, whether the child was in school during the time use survey, the year the child started school, and the child dependency ratio of the household.

TABLE 4: WEALTH QUARTILES

Math scores and share of time spent on play

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent on play	2.389**	1.615	-0.440	-8.028
	(0.895)	(1.238)	(1.384)	(5.579)
Constant	-0.0828	-1.059	0.0591	2.004*
	(0.480)	(0.769)	(0.653)	(1.060)
Observations	446	455	459	445
R-squared	0.170	0.240	0.200	0.165
Demographic controls	YES	YES	YES	YES
Wealth controls	YES	YES	YES	YES
Interactions	YES	YES	YES	YES

Reading scores and share of time spent on play

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent on play	1.054	-1.342	2.262	0.0622
	(1.099)	(1.381)	(1.339)	(3.200)
Constant	2.278**	0.867	-0.355	1.388**
	(0.971)	(1.025)	(0.456)	(0.567)
Observations	429	443	455	441
R-squared	0.228	0.211	0.180	0.164
Demographic controls	YES	YES	YES	YES
Wealth controls	YES	YES	YES	YES
Interactions	YES	YES	YES	YES

Vocabulary scores and share of time spent on play

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent on play	1.405*	-1.322	1.266	-0.864
	(0.691)	(1.188)	(1.001)	(3.102)
Constant	0.381	0.505	-0.640	-0.0150
	(0.807)	(0.604)	(0.432)	(0.607)
Observations	436	447	454	443
R-squared	0.266	0.321	0.269	0.155
Demographic controls	YES	YES	YES	YES
Wealth controls	YES	YES	YES	YES
Interactions	YES	YES	YES	YES

Educational attainment and share of time spent on play

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent on play	0.390	-3.113**	-1.657*	-2.868
	(0.880)	(1.450)	(0.817)	(1.832)
Constant	15.12***	14.13***	12.54***	12.41***
	(0.869)	(0.864)	(0.626)	(0.847)
Observations	419	443	453	438
R-squared	0.347	0.253	0.329	0.262
Demographic controls	YES	YES	YES	YES
Wealth controls	YES	YES	YES	YES
Interactions	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note. All estimates were obtained using a linear probability model. Standard errors are clustered by sentinel site. Controls include gender, indigenous first language, indigenous ethnicity, primary caregiver education level, owning a working television, whether the child was in school during the time use survey, the year the child started school, and the child dependency ratio of the household.

TABLE 5: RURAL AND URBAN OUTCOMES

Share of time spent on play and educational outcomes, Rural

VARIABLES	(1) Math z-score	(2) Reading z-score	(3) Vocab z-score	(4) Grade in 2016
Share of time spent on play	0.550	-1.344	0.179	-0.867
	(0.598)	(1.120)	(0.462)	(1.021)
Constant	-0.500	0.276	-0.0565	15.20***
	(0.638)	(1.037)	(1.063)	(0.865)
Observations	557	533	541	533
R-squared	0.138	0.207	0.286	0.293
Demographic controls	YES	YES	YES	YES
Wealth controls	YES	YES	YES	YES
Interactions	YES	YES	YES	YES

Share of time spent on play and educational outcomes, Urban

VARIABLES	(1) Math z-score	(2) Reading z-score	(3) Vocab z-score	(4) Grade in 2016
Share of time spent on play	2.107*	0.852	0.956	-1.343*
	(1.087)	(1.131)	(0.849)	(0.694)
Constant	-1.080**	-0.148	-0.483	12.15***
	(0.466)	(0.428)	(0.321)	(0.531)
Observations	1,248	1,235	1,239	1,220
R-squared	0.184	0.125	0.245	0.237
Demographic controls	YES	YES	YES	YES
Wealth controls	YES	YES	YES	YES
Interactions	YES	YES	YES	YES

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note. All estimates were obtained using a linear probability model. Standard errors are clustered by sentinel site. Controls include wealth, gender, indigenous first language, indigenous ethnicity, primary caregiver education level, owning a working television, whether the child was in school during the time use survey, the year the child started school, and the child dependency ratio of the household.

Table 6: Determinants of attrition

VARIABLES	(1) Ever missed a round	(2) Not in round 2	(3) Not in round 5
Wealth index	-0.271	-0.501*	-0.286
	(0.183)	(0.262)	(0.199)
Male	-0.0576	-0.147	-0.0797
	(0.0725)	(0.103)	(0.0778)
Caregiver's first language is indigenous	-0.0641	-0.0907	0.00837
	(0.115)	(0.180)	(0.123)
Indigenous ethnicity	-0.104		-0.170
	(0.111)		(0.122)
Caregiver completed secondary education, round 1	-0.0511	-0.177	0.0501
	(0.0910)	(0.128)	(0.0975)
Household size	-0.00429	-0.0199	0.00509
	(0.0157)	(0.0229)	(0.0167)
Any events that decreased household welfare since pregnancy with child	-0.0169	0.0235	-0.00898
	(0.0746)	(0.104)	(0.0800)
Do you feel part of the community	-0.235**	-0.0539	-0.210**
	(0.0938)	(0.138)	(0.101)
Constant	-0.768***	-1.142***	-1.009***
	(0.153)	(0.217)	(0.164)
Observations	2,035	1,723	2,035
Pseudo R-squared	0.00745	0.0189	0.00705

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note. The dependent variable of interest is listed below the column number. All estimates were obtained using a probit model. A higher wealth index score indicates more wealth. All of the children with an indigenous ethnicity were observed in survey round two.

Table 7: Correlation between time use categories

	Play	Sleep	School	Study	Chores	Caring for others	Unpaid work
Play	1						
Sleep	-0.0933***	1					
School	-0.666***	-0.429***	1				
Study	-0.482***	-0.356***	0.513***	1			
Chores	-0.180***	0.00852	-0.182***	-0.168***	1		
Caring for others	-0.137***	-0.201***	-0.124***	-0.122***	0.144***	1	
Unpaid work	-0.0935***	-0.0137	-0.159***	-0.143***	0.123***	0.0281	1

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

Table 8: Share of time spent on play and educational outcomes

VARIABLES	(1) Math z-score	(2) Reading z-score	(3) Vocab z-score	(4) Grade in 2016
Share of time spent on play	1.273	0.199	0.874	-1.039
	(0.888)	(0.970)	(0.557)	(0.635)
Wealth index	0.598**	0.488***	0.362**	0.337**
	(0.221)	(0.165)	(0.169)	(0.160)
Male	0.184**	-0.225**	0.0337	-0.166*
	(0.0768)	(0.104)	(0.0806)	(0.0934)
Child's first language is indigenous	-0.0370	-0.500	-0.560**	-0.257
	(0.200)	(0.295)	(0.223)	(0.274)
Indigenous ethnicity	-0.0461	-0.111	-0.0941	0.0872
	(0.170)	(0.180)	(0.193)	(0.119)
Rural	-0.0201	-0.0860	-0.0527	-0.115
	(0.187)	(0.169)	(0.164)	(0.0754)
Had a sibling ages 4-10	-0.206**	-0.111	-0.0202	-0.108
	(0.0949)	(0.115)	(0.0925)	(0.111)
Caregiver completed secondary education	0.300**	0.208*	0.377***	0.215*
	(0.138)	(0.101)	(0.0786)	(0.109)
Household owns working television	0.146	0.221*	0.295***	0.0105
	(0.106)	(0.110)	(0.0740)	(0.0921)
In school during time use survey	0.0723	0.0774	0.261***	0.159
	(0.0745)	(0.0983)	(0.0864)	(0.0926)
Year started school	-0.0728***	-0.148***	-0.0828**	-0.498***
	(0.0216)	(0.0366)	(0.0346)	(0.0714)
Child dependency ratio	0.141**	0.00826	-0.0236	-0.102
	(0.0590)	(0.0997)	(0.0863)	(0.0623)
Proportion of school years in private schools	0.315**	0.280**	0.323***	0.0840
	(0.117)	(0.123)	(0.0610)	(0.0699)
Proportion of school years in parochial or NGO schools	0.393	-0.0147	-0.0144	-1.923*
	(0.485)	(0.401)	(0.311)	(0.950)
Proportion of school years in other schools	0.485	-0.548	-0.0354	0.0937
	(0.453)	(0.534)	(0.546)	(0.987)
Share of time spent playing x Child dependency ratio	-0.598***	-0.235	0.112	-0.0129
	(0.191)	(0.401)	(0.241)	(0.240)

Table 9: Share of time spent on play and math test scores, by wealth quartile

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent on play	2.389**	1.615	-0.440	-8.028
	(0.895)	(1.238)	(1.384)	(5.579)
Male	0.125	0.331*	0.187	-0.00326
	(0.158)	(0.159)	(0.210)	(0.166)
Child's first language is indigenous	-0.119	0.134	-0.640	1.492***
	(0.257)	(0.386)	(0.664)	(0.0966)
Indigenous ethnicity	-0.00886	0.422	-0.437	-0.444
	(0.166)	(0.341)	(0.445)	(0.658)
Rural	-0.0457	-0.467	0.986**	0.0829
	(0.272)	(0.317)	(0.446)	(0.340)
Had a sibling ages 4-10	-0.136	0.0280	-0.485**	-0.133
	(0.198)	(0.237)	(0.199)	(0.294)
Caregiver completed secondary education	1.010*	0.663*	-0.0389	0.222
	(0.496)	(0.347)	(0.251)	(0.283)
Household owns working television	0.213	0.206	0.361	-1.869
	(0.256)	(0.187)	(0.281)	(1.312)
In school during time use survey	0.124*	0.106	-0.124	0.136
	(0.0663)	(0.149)	(0.220)	(0.264)
Year started school	-0.138**	-0.00695	-0.0449	-0.123
	(0.0522)	(0.0663)	(0.0522)	(0.0719)
Child dependency ratio	0.144	0.109	-0.0265	0.355
	(0.102)	(0.0951)	(0.124)	(0.247)
Proportion of school years in private schools	-0.0797	0.973**	0.433	0.221
	(0.542)	(0.425)	(0.267)	(0.167)
Proportion of school years in parochial or NGO schools	0.292	2.595**	1.832*	-1.088
	(2.332)	(0.979)	(0.933)	(0.718)
Proportion of school years in other schools	0.390	0.231	0.421	1.330
	(0.493)	(0.918)	(0.534)	(0.924)
Share of time spent playing x Child dependency ratio	-0.679	-0.395	0.358	-1.832
	(0.399)	(0.371)	(0.388)	(1.420)
Share of time spent playing x Male	0.169	-0.208	0.469	0.808
	(0.626)	(0.495)	(0.870)	(0.985)

Table 9, continued

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent playing x Indigenous ethnicity	-0.605	-1.826	1.290	3.648*
	(0.659)	(1.425)	(1.430)	(2.080)
Share of time spent playing x Had sibling age 4-10	0.0300	-0.356	2.024***	-0.350
	(0.787)	(0.885)	(0.695)	(1.473)
Share of time spent playing x Caregiver completed secondary education	-2.364	-0.724	2.047*	1.231
	(1.640)	(1.325)	(0.995)	(1.209)
Share of time spent playing x Household owns working television	-0.612	0.160	-2.313**	9.048
	(0.904)	(0.693)	(0.958)	(5.470)
Share of time spent playing x Indigenous first language	-0.459	0.427	5.312***	
	(0.463)	(0.966)	(1.338)	
Share of time spent playing x Rural	-0.500	0.972	-3.617**	
	(0.739)	(0.953)	(1.584)	
Constant	-0.0828	-1.059	0.0591	2.004*
	(0.480)	(0.769)	(0.653)	(1.060)
Observations	446	455	459	445
R-squared	0.170	0.240	0.200	0.165

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note. All estimates were obtained using a linear probability model. Standard errors are clustered by sentinel site. In the fourth wealth quantile, there is only one rural child, so the interaction term is perfectly collinear with rural. Similarly, in the fourth wealth quantile, there are only four children with an indigenous first language, and the interaction term is perfectly collinear with indigenous first language. Finally, the fourth wealth quantile also only had one child from one of the sentinel sites (cluster 16), and this child had an indigenous first language, and thus there was perfect collinearity between the cluster and indigenous first language.

Table 10: Share of time spent on play and reading scores, by wealth quartile

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent on play	1.054	-1.342	2.262	0.0622
	(1.099)	(1.381)	(1.339)	(3.200)
Male	-0.0709	-0.261	-0.158	-0.333
	(0.221)	(0.173)	(0.240)	(0.226)
Child's first language is indigenous	-0.231	-0.710***	-1.523**	0.193*
	(0.389)	(0.231)	(0.582)	(0.101)
Indigenous ethnicity	-0.264	0.452**	-0.340	-0.540
	(0.285)	(0.163)	(0.210)	(0.479)
Rural	-0.331*	-0.0890	0.206	-1.740***
	(0.182)	(0.206)	(0.589)	(0.358)
Had a sibling ages 4-10	-0.00436	-0.211	-0.470*	0.0526
	(0.219)	(0.211)	(0.253)	(0.290)
Caregiver completed secondary education	1.121**	0.0637	0.163	0.204
	(0.531)	(0.289)	(0.254)	(0.188)
Household owns working television	0.0270	0.302	0.822**	-0.911
	(0.300)	(0.222)	(0.329)	(0.814)
In school during time use survey	0.0583	0.0492	0.153	-0.297
	(0.0980)	(0.254)	(0.182)	(0.187)
Year started school	-0.333**	-0.202**	-0.124***	-0.0595
	(0.129)	(0.0955)	(0.0400)	(0.0546)
Child dependency ratio	-0.126	-0.0782	0.220**	0.548**
	(0.156)	(0.0979)	(0.104)	(0.244)
Proportion of school years in private schools	-0.897	-0.00643	0.379**	0.366*
	(1.209)	(0.346)	(0.179)	(0.183)
Proportion of school years in parochial or NGO schools	6.327*	1.060	0.711	-0.201
	(3.312)	(0.631)	(0.786)	(0.688)
Proportion of school years in other schools	-3.735***	0.749	0.931	-0.797
	(0.282)	(1.296)	(0.894)	(0.787)
Share of time spent playing x Child dependency ratio	-0.149	0.310	-0.587	-2.802*
	(0.447)	(0.328)	(0.621)	(1.373)
Share of time spent playing x Male	0.484	1.695**	0.730	1.039
	(0.840)	(0.703)	(1.052)	(1.284)

Table 10, continued

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent playing x Indigenous ethnicity	0.880	-2.276***	0.854	3.023
	(0.887)	(0.644)	(0.780)	(1.761)
Share of time spent playing x Had sibling age 4-10	-0.859	0.667	1.454*	-0.349
	(0.931)	(0.955)	(0.840)	(1.546)
Share of time spent playing x Caregiver completed secondary education	-3.504*	1.161	1.013	0.653
	(1.776)	(1.399)	(1.023)	(0.957)
Share of time spent playing x Household owns working television	-0.0603	0.129	-3.021**	0.843
	(1.126)	(0.870)	(1.236)	(3.264)
Share of time spent playing x Indigenous first language	-0.391	2.964***	5.093**	
	(0.989)	(0.958)	(1.830)	
Share of time spent playing x Rural	-0.204	-0.852	0.164	
	(0.569)	(0.824)	(2.322)	
Constant	2.278**	0.867	-0.355	1.388**
	(0.971)	(1.025)	(0.456)	(0.567)
Observations	429	443	455	441
R-squared	0.228	0.211	0.180	0.164

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note. All estimates were obtained using a linear probability model. Standard errors clustered by sentinel site. In the fourth wealth quantile, there is only one rural child, so the interaction term is perfectly collinear with rural. Similarly, in the fourth wealth quantile, there are only four children with an indigenous first language, and the interaction term is perfectly collinear with indigenous first language. Finally, the fourth wealth quantile also only had one child from one of the sentinel sites (cluster 16), and this child had an indigenous first language, and thus there was perfect collinearity between the cluster and indigenous first language.

Table 11: Share of time spent on play and vocabulary scores, by wealth quartile

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent on play	1.405*	-1.322	1.266	-0.864
	(0.691)	(1.188)	(1.001)	(3.102)
Male	0.198	-0.0742	0.0251	0.209
	(0.170)	(0.156)	(0.159)	(0.142)
Child's first language is indigenous	-0.260	-0.800**	-0.452	-0.0759
	(0.185)	(0.372)	(0.500)	(0.122)
Indigenous ethnicity	-0.316**	0.609*	-0.497**	-0.495
	(0.149)	(0.329)	(0.186)	(0.474)
Rural	-0.0574	-0.573**	0.964**	-1.011***
	(0.223)	(0.242)	(0.458)	(0.203)
Had a sibling ages 4-10	0.119	-0.173	-0.258	-0.236
	(0.220)	(0.182)	(0.170)	(0.176)
Caregiver completed secondary education	0.772***	0.255	0.354**	0.836***
	(0.245)	(0.275)	(0.147)	(0.271)
Household owns working television	0.626***	0.144	0.388	-0.406
	(0.150)	(0.183)	(0.233)	(0.545)
In school during time use survey	0.268**	0.219**	0.304*	0.112
	(0.127)	(0.104)	(0.156)	(0.315)
Year started school	-0.185	-0.0840	-0.0591	-0.0512
	(0.114)	(0.0692)	(0.0423)	(0.0779)
Child dependency ratio	-0.151	-0.0604	0.0864	0.255
	(0.141)	(0.141)	(0.122)	(0.231)
Proportion of school years in private schools	0.877	0.813***	0.374*	0.350***
	(0.541)	(0.253)	(0.186)	(0.0837)
Proportion of school years in parochial or NGO schools	2.168	0.593	-0.666	0.472
	(2.614)	(0.631)	(0.548)	(0.490)
Proportion of school years in other schools	-1.218*	1.812	0.593	0.252
	(0.592)	(1.499)	(1.064)	(0.946)
Share of time spent playing x Child dependency ratio	0.464	0.214	-0.0462	-1.072
	(0.419)	(0.432)	(0.475)	(0.964)
Share of time spent playing x Male	-0.212	1.628**	0.831	-0.961
	(0.659)	(0.634)	(0.725)	(0.637)

Table 11, continued

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent playing x Indigenous ethnicity	1.003*	-2.931**	2.297**	2.452
	(0.566)	(1.138)	(0.922)	(1.566)
Share of time spent playing x Had sibling age 4-10	-1.311*	-0.130	0.143	0.347
	(0.752)	(0.718)	(0.811)	(0.818)
Share of time spent playing x Caregiver completed secondary education	-1.699*	0.330	0.0400	-2.471*
	(0.882)	(1.112)	(0.545)	(1.238)
Share of time spent playing x Household owns working television	-2.019***	0.0772	-1.632*	4.452*
	(0.610)	(0.763)	(0.902)	(2.278)
Share of time spent playing x Indigenous first language	-0.589	1.890*	2.100	
	(0.512)	(0.932)	(1.451)	
Share of time spent playing x Rural	-0.510	0.617	-3.581	
	(0.704)	(0.966)	(2.721)	
Constant	0.381	0.505	-0.640	-0.0150
	(0.807)	(0.604)	(0.432)	(0.607)
Observations	436	447	454	443
R-squared	0.266	0.321	0.269	0.155

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note. All estimates were obtained using a linear probability model. Standard errors clustered by sentinel site. In the fourth wealth quantile, there is only one rural child, so the interaction term is perfectly collinear with rural. Similarly, in the fourth wealth quantile, there are only four children with an indigenous first language, and the interaction term is perfectly collinear with indigenous first language. Finally, the fourth wealth quantile also only had one child from one of the sentinel sites (cluster 16), and this child had an indigenous first language, and thus there was perfect collinearity between the cluster and indigenous first language.

Table 12: Share of time spent on play and educational attainment, by wealth quartile

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent on play	0.390	-3.113**	-1.657*	-2.868
	(0.880)	(1.450)	(0.817)	(1.832)
Male	-0.284	-0.0477	-0.205	-0.213
	(0.175)	(0.204)	(0.192)	(0.171)
Child's first language is indigenous	-0.354	0.331*	-0.719	0.220
	(0.421)	(0.177)	(0.470)	(0.136)
Indigenous ethnicity	-0.0445	0.378	0.164	-0.273
	(0.198)	(0.249)	(0.168)	(0.348)
Rural	-0.0683	-0.369*	0.170	0.131
	(0.151)	(0.208)	(0.338)	(0.328)
Had a sibling ages 4-10	-0.0702	-0.301	-0.295	0.222
	(0.204)	(0.274)	(0.238)	(0.189)
Caregiver completed secondary education	0.348	0.0279	-0.0634	0.579**
	(0.488)	(0.340)	(0.113)	(0.204)
Household owns working television	-0.0486	-0.0235	-0.0634	-0.544***
	(0.306)	(0.207)	(0.139)	(0.167)
In school during time use survey	0.327***	-0.105	0.0873	-0.00642
	(0.103)	(0.192)	(0.145)	(0.350)
Year started school	-0.852***	-0.620***	-0.411***	-0.341***
	(0.118)	(0.110)	(0.0856)	(0.0865)
Child dependency ratio	-0.0447	-0.176	-0.113	-0.124
	(0.0814)	(0.202)	(0.133)	(0.213)
Proportion of school years in private schools	0.165	0.577	0.275***	0.0485
	(0.285)	(0.355)	(0.0958)	(0.0606)
Proportion of school years in parochial or NGO schools	-4.922**	-1.339**	-0.0189	-2.216
	(2.187)	(0.540)	(0.435)	(1.382)
Proportion of school years in other schools	-1.174***	1.680*	2.330***	-1.555
	(0.405)	(0.939)	(0.748)	(1.343)
Share of time spent playing x Child dependency ratio	-0.314	0.0674	-0.113	0.537
	(0.514)	(0.708)	(0.388)	(1.263)
Share of time spent playing x Male	1.075	0.202	0.523	0.720
	(0.725)	(0.545)	(0.852)	(1.134)

Table 12, continued

VARIABLES	(1) Q1	(2) Q2	(3) Q3	(4) Q4
Share of time spent playing x Indigenous ethnicity	0.527	-1.124	0.570	1.082
	(0.830)	(1.333)	(0.789)	(1.304)
Share of time spent playing x Had sibling age 4-10	0.198	1.602*	1.116	-1.680
	(0.689)	(0.926)	(1.051)	(0.969)
Share of time spent playing x Caregiver completed secondary education	0.351	0.842	1.233**	-1.216
	(2.204)	(1.119)	(0.518)	(1.467)
Share of time spent playing x Household owns working television	-0.431	1.404	-0.326	2.038**
	(1.006)	(0.884)	(0.687)	(0.951)
Share of time spent playing x Indigenous first language	0.735	-1.661*	1.687	
	(0.870)	(0.835)	(1.408)	
Share of time spent playing x Rural	-0.794	1.490**	-0.730	
	(0.609)	(0.682)	(1.310)	
Constant	15.12***	14.13***	12.54***	12.41***
	(0.869)	(0.864)	(0.626)	(0.847)
Observations	419	443	453	438
R-squared	0.347	0.253	0.329	0.262

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note. All estimates were obtained using a linear probability model. Standard errors clustered by sentinel site. In the fourth wealth quantile, there is only one rural child, so the interaction term is perfectly collinear with rural. Similarly, in the fourth wealth quantile, there are only four children with an indigenous first language, and the interaction term is perfectly collinear with indigenous first language. Finally, the fourth wealth quantile also only had one child from one of the sentinel sites (cluster 16), and this child had an indigenous first language, and thus there was perfect collinearity between the cluster and indigenous first language.

Table 13: Share of time spent on play educational outcomes, Rural

	(1)	(2)	(3)	(4)
VARIABLES	Math z-score	Reading z-score	Vocab z-score	Grade in 2016
Share of time spent on play	0.550	-1.344	0.179	-0.867
	(0.598)	(1.120)	(0.462)	(1.021)
Wealth index	0.826**	1.252***	0.674**	0.929**
	(0.371)	(0.347)	(0.305)	(0.389)
Male	0.0942	-0.155	0.144	-0.205
	(0.143)	(0.215)	(0.102)	(0.146)
Child's first language is indigenous	-0.0674	-0.440	-0.531***	-0.201
	(0.215)	(0.286)	(0.173)	(0.315)
Indigenous ethnicity	-0.000314	-0.116	0.00544	0.207
	(0.176)	(0.236)	(0.271)	(0.135)
Had a sibling ages 4-10	-0.0311	-0.121	0.0963	-0.226
	(0.169)	(0.182)	(0.190)	(0.148)
Caregiver completed secondary education	0.744	0.687	0.379	-0.381
	(0.465)	(0.487)	(0.219)	(0.489)
Household owns working television	-0.220	0.0636	0.276**	-0.179
	(0.146)	(0.208)	(0.107)	(0.155)
In school during time use survey	0.0239	-0.0926	0.272	0.132
	(0.108)	(0.135)	(0.165)	(0.153)
Year started school	-0.0950	-0.261**	-0.178	-0.790***
	(0.0668)	(0.100)	(0.115)	(0.108)
Child dependency ratio	0.0170	-0.173	-0.183	-0.0744
	(0.0994)	(0.141)	(0.106)	(0.0611)
Proportion of school years in private schools	0.891***	0.179	1.191***	0.269
	(0.282)	(0.583)	(0.289)	(0.433)
Proportion of school years in parochial or NGO schools	1.336	-1.278	5.496***	1.819
	(1.601)	(2.105)	(0.949)	(1.974)
Proportion of school years in other schools	0.107	-2.519***	-1.302***	-0.994***
	(0.263)	(0.197)	(0.211)	(0.295)
Share of time spent playing x Child dependency ratio	-0.261	0.193	0.712**	-0.313
	(0.429)	(0.481)	(0.300)	(0.406)
Share of time spent playing x Male	0.250	1.158	0.235	0.846
	(0.630)	(0.802)	(0.403)	(0.566)

Table 13, continued

VARIABLES	(1) Math z-score	(2) Reading z-score	(3) Vocab z-score	(4) Grade in 2016
Share of time spent playing x Indigenous first language	-0.165	1.045	0.286	0.325
	(0.303)	(0.807)	(0.621)	(0.697)
Share of time spent playing x Indigenous ethnicity	-0.697	-0.0776	-0.195	-0.452
	(0.643)	(0.813)	(0.926)	(0.417)
Share of time spent playing x Had sibling age 4-10	0.120	-0.257	-1.441**	0.548
	(0.713)	(0.552)	(0.671)	(0.716)
Share of time spent playing x Caregiver completed secondary education	-1.464	-0.963	-0.402	3.315*
	(1.569)	(1.491)	(1.055)	(1.633)
Share of time spent playing x Household owns working television	1.328	0.111	-0.362	0.764
	(0.840)	(1.004)	(0.569)	(0.595)
Constant	-0.500	0.276	-0.0565	15.20***
	(0.638)	(1.037)	(1.063)	(0.865)
Observations	557	533	541	533
R-squared	0.138	0.207	0.286	0.293

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note. All estimates were obtained using a linear probability model. Standard errors clustered by sentinel site. A higher wealth index score indicates more wealth.

Table 14: Share of time spent on play and educational outcomes, Urban

VARIABLES	(1) Math z-score	(2) Reading z-score	(3) Vocab z-score	(4) Grade in 2016
Share of time spent on play	2.107*	0.852	0.956	-1.343*
	(1.087)	(1.131)	(0.849)	(0.694)
Wealth index	0.546**	0.326	0.338**	0.235
	(0.214)	(0.197)	(0.159)	(0.165)
Male	0.212**	-0.266***	0.0281	-0.189
	(0.0955)	(0.0910)	(0.0920)	(0.118)
Child's first language is indigenous	0.554	-0.382	-0.345	-0.437*
	(0.387)	(0.355)	(0.268)	(0.213)
Indigenous ethnicity	-0.0563	-0.154	-0.158	-0.151
	(0.335)	(0.308)	(0.240)	(0.141)
Had a sibling ages 4-10	-0.312***	-0.197	-0.157*	-0.0898
	(0.0902)	(0.137)	(0.0833)	(0.142)
Caregiver completed secondary education	0.245	0.195*	0.374***	0.235**
	(0.147)	(0.0952)	(0.0782)	(0.0982)
Household owns working television	0.384*	0.259*	0.143	0.0526
	(0.184)	(0.147)	(0.123)	(0.148)
In school during time use survey	0.119	0.193*	0.196*	0.130
	(0.0944)	(0.111)	(0.0938)	(0.107)
Year started school	-0.0606*	-0.108***	-0.0534	-0.423***
	(0.0300)	(0.0335)	(0.0326)	(0.0655)
Child dependency ratio	0.264**	0.219**	0.178*	-0.0907
	(0.0968)	(0.0956)	(0.102)	(0.109)
Proportion of school years in private schools	0.325**	0.346**	0.354***	0.148**
	(0.125)	(0.127)	(0.0633)	(0.0644)
Proportion of school years in parochial or NGO schools	0.467	0.129	0.0605	-1.628*
	(0.483)	(0.424)	(0.343)	(0.916)
Proportion of school years in other schools	0.683	-0.0273	0.656	0.394
	(0.559)	(0.521)	(0.492)	(1.183)
Share of time spent playing x Child dependency ratio	-0.907**	-0.723	-0.579	0.0629
	(0.398)	(0.478)	(0.369)	(0.351)
Share of time spent playing x Male	0.133	1.011**	0.572	0.424
	(0.352)	(0.423)	(0.460)	(0.526)

Table 14, continued

VARIABLES	(1) Math z-score	(2) Reading z-score	(3) Vocab z-score	(4) Grade in 2016
Share of time spent playing x Indigenous first language	-0.956	1.016	1.293	1.121
	(1.701)	(1.612)	(1.521)	(1.091)
Share of time spent playing x Indigenous ethnicity	1.121	0.761	0.952	1.089**
	(1.079)	(1.231)	(1.136)	(0.514)
Share of time spent playing x Had sibling age 4-10	0.696	0.491	-0.219	0.372
	(0.451)	(0.767)	(0.377)	(0.584)
Share of time spent playing x Caregiver completed secondary education	0.654	0.335	-0.208	-0.0959
	(0.703)	(0.475)	(0.343)	(0.443)
Share of time spent playing x Household owns working television	-1.397**	-0.666	-0.512	0.196
	(0.583)	(0.623)	(0.392)	(0.536)
Constant	-1.080**	-0.148	-0.483	12.15***
	(0.466)	(0.428)	(0.321)	(0.531)
Observations	1,248	1,235	1,239	1,220
R-squared	0.184	0.125	0.245	0.237

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note. All estimates were obtained using a linear probability model. Standard errors clustered by sentinel site. A higher wealth index score indicates more wealth.

The King and the Cronies: An Institutional Analysis of the Philippines During the Marcos Regime

Paul Gabriel L. Cosme

June 2020

1 Introduction

The Philippines is an archipelago endowed with natural resources, from crops to precious metals and minerals. However, its institutions are ridden with extractive elements that prevent Filipinos from flourishing. This is the reason why the Philippines is called an Oligarchipelago (Almeida, 2012). A fundamental question arises then: when and how did the Philippines succumb to such extraction? An instinctive answer lies in the regime of Ferdinand Marcos, the Filipino authoritarian president from 1965 to 1986. Scholars argue (Hutchcroft, 1991) that Marcos disrupted the political culture that can be described as patrimonial.[1] But to call Marcos an aberration is misleading. As Filipino historian Nicole CuUnjieng Aboitiz (2009) notes, Marcos is not an aberration to the culture; he is its apotheosis.

The culture of patrimonialism provides the junctures to extract economic resources, and Marcos exploited this culture to further his and his cronies' gains—severely damaging polit-

[1] See note on Max Weber's patrimonialism in Almeida (2012).

ical and economic institutions. Understanding the origins of such extractive institutions is key to developing solutions to achieve an open-access order society (North et al., 2009). In this exploration, I survey the economic and political landscape of the Philippines during the Pre-Marcos era. I touch on colonial influences of Spain and the United States on Filipino institutions. Then, I move to economic and political developments after the Philippines gained independence from the US in 1946. Afterwards, I focus on giving a background on political and economic institutions that Marcos placed during his regime. I proceed with my analysis of both political and economic institutions. Finally, I conclude by reflecting how these institutions still plague today's Philippines.

2 History and Background

The Philippines is composed of about 7,000 islands, 1,000 of which are inhabited (Dohner et al., 1989). Being a former colony of both Spain and the US, each colonizer played a vital role in molding the country, especially in forming the culture of patrimonialism, clientelism, and oligarchy.

2.1 Pre-Marcos Era

During the Spanish colonial period, the Spaniards divided the Philippines into estates called *encomiendas*. They gave these estates to Spanish settlers and the church which allowed them to receive tribute from natives. They were not land grants but a means to support the settlers. Then, during the Industrial Revolution, the surge in demand for agricultural production resulted into the Royal Decree of February 13, 1894. This decree gave landholders a year to secure legal titles to their land. This resulted into the expansion of estates and misappropriation of lands of least 400,000 people. Many small landowners were unaware of the decree, and their lands were included in the titles of big landowners. These huge estates were called *haciendas*, and even though they were fruits of capitalist development, they were still feudal in nature as absentee landlords in Manila dealt with commercial export (Litonjua, 2001). Afterwards, the Philippines became a US colony (until 1946) as a result of the Spanish-American War. The US intensely molded the language, education system, and institutions in the Philippines, but it never changed the hacienda system from the Spanish colonial period (Dohner et al., 1989). This land system gave rise to the ilustrados (educated). This new class composed the economic and political power elite in the Philippines then and until now (Litonjua, 2001).

As a classic colonial state, the Spanish relied on native allies to govern as collaborators. These allies were selected through a restricted electoral process (Litonjua, 2001). The Spanish saw political office as the right to serve, but Filipinos saw this as "surrogates for struggles to survive, to manipulate patron-client relations, to advance economically, and to wield a measure power in a situation in which they were inherently subordinate and subservient." (Litonjua, 2001). Instead of honing an ethos of public service, Filipino bureaucrats learned to manipulate and exploit the government. This mentality deepened under US colonialism and persists today. This resulted in "a small landed elite who held economic power, and a weak and incompetent bureaucracy mired in the politics of patronage and clientelism." (Litonjua, 2001).

After independence, these patrimonial features strengthened through three factors: (1) there was a blurring of the distinction between official and private spheres as personal contacts became more needed to enter the central bureaucracy (Hutchcroft, 1991). Moreover, (2) clientelism underwent significant changes. Patrons expanded opportunities by gaining access to external and "office-based" state resources. Local patrons and oligarchs did not lose power in relation to the state, but the role of state resources increased. These oligarchs gained "enormous power to milk the central state's 'particularistic distributive capacity' " (Hutchcroft, 1991). Finally, (3) oligarchs diversified their interests beyond agriculture to include commerce, manufacturing, finance, and others. This made access to the state machinery more important. Therefore, "as long as 'rents' can be obtained[...]rent-seekers find it more important to maintain government connections than to concern themselves with the 'internal efficiencies and investments' of their firms." (Hutchcroft, 1991). These developments intensified the role of state in strengthening patrimonial features and securing patronage and rents that are seen today.

2.2 Marcos Era

The Philippine Republic until 1972 can be accurately described as an elite democracy (Litonjua, 2001). Sixty dynastic families "whose wealth originally emerged during the Spanish colonial era[...]constitute one-fifth of the population but receive half of the country's income." (Litonjua, 2001) These families inherited the state control from the US, and they used it to maintain their hegemonic foothold in the country (Litonjua, 2001). These Filipino political families influenced and participated in politics and elections through thuggery, violence, and money, that the Filipino society was described as "an anarchy of families." (Litonjua, 2001)

By the late 1960s, the elite democracy was shaken. The economy was stagnant. Communists were revitalizing. There was civil unrest among Muslims in the south. The violence of political warlords or violence specialists (North et al., 2009) was pervasive, and student activism was in full blow. Marcos exacerbated this crisis as he became the first president to be reelected in 1969. He took advantage of this crisis to declare martial law. Scholars claim that Marcos broke the political tradition upon declaring martial law (CuUnjieng Aboitiz, 2009), and that he disrupted a patrimonial system that ran within Filipino dynasties. Members of these dynasties are addressed as the bosses in their local communities. Marcos managed to concentrate power in his hands and became the single predatory boss.[2] However, CuUnjieng Aboitiz (2009) contends that Marcos was the apotheosis of the political culture as he made the political system work for him. He left the courts untouched; however, the courts unwittingly legalized Marcos's authoritarian regime.

Marcos gradually stripped political participation due to a series of events. In the middle of the constitutional convention in 1971, a plaza bombing impelled Marcos to suspend the writ of habeas corpus. He argued that the bombing was a Communist terrorist plot. There were petitions to the Supreme Court questioning the validity of this suspension, but the court upheld the suspension affirming the Communist threat. Marcos left courts to function healthily to unwittingly legitimize his authoritarian regime; thus, removing all obstacles for his upcoming declaration of martial law a year later (CuUnjieng Aboitiz, 2009). Cesar Virata, Marcos's prime minister, justified martial law by saying that the Philippines's neighbors were tightening control and security, and that Marcos was just following them (CuUnjieng Aboitiz, 2009). However, this declaration was a step to solidify Marcos's economic and political power concentration among the elite families.

[2]See Sidel's *Bossism* in CuUnjieng Aboitiz (2009), "Ferdinand Marcos: 'Apotheosis' of the Philippine Historical Political Tradition," p. 3.

Marcos has profound influence on the composition of the elites and cronies in the country. Cronies are not the same as oligarchs. Cronies "are particularly favored by the current regime, regardless of their origins" (Hutchcroft, 1991), while an oligarch "may not be a current crony but in either case has already established [their] fortune in earlier dispositions."[3] In a patrimonial style, Marcos targeted those rival families and oligarchs who threatened his clan (Hutchcroft, 1991). Thus, using the state, Marcos skillfully manipulated the political culture to maximize his cronies' and his own gains. He legalized his relatives' and cronies' monopolies and allowed them to receive kickbacks as well (Dohner et al., 1989). Marcos relied on expansionary fiscal and monetary policies to benefit his cronies, but local and foreign debt grew consequently (Dohner et al., 1989). This debt heavily crippled the economy and persists until now.

3 Institutional Analysis

3.1 Political Institutions

During the Marcos regime, the greatest blowback to political institutions, perhaps, is the curtailing of political participation. Marcos committed multiple human rights violations and suppressed press freedom. American historian Alfred Mccoy (2001) argues that Marcos created a culture of fear in the Philippines by making protesters and political enemies disappear. Political scientist David A. Mccoy (2001) also notes that when Marcos declared martial law, the first major action he took was to shut down every major newspaper, radio, and television station. Only one newspaper, one television station, and the government-

[3] See "The Rise of the Cronies and the Relatives" in Hutchcroft (1991), p. 434.

owned radio station can operate within the country. The military enforced this restriction, and a Marcos-controlled committee monopolized the media industry.

Suppressing press freedom limited public awareness about government atrocities. Because of this, only a few people felt the need to organize against the administration. Some Filipinos even lauded Marcos for allegedly keeping the country safe from any Communist threat. Manipulating the press created an illusion of security within the nation. Marcos could sustain his political power concentration by suppressing press freedom, and consequently, political participation. Even the elite dynastic families are afraid to counteract Marcos, because he "controlled the levers of power and coercion, he monopolized the dispensation of patronage and privilege, his was a personal rule without ideological and institutional constraints" (Litonjua, 2001). Nevertheless, he needed the support of his cronies and the military to control key sectors in the economy and the state. If they defect, Marcos loses power. In this sense, Marcos, as the single predatory and patrimonial boss, effectively executed political extraction to reward rents to his cronies and allies. This only shows that Marcos was not an aberration to the political culture as scholars depict him—he was its apotheosis.

3.2 Economic Institutions

Crony capitalism is the best way to describe Marcos's economic policies and institutions. By concentrating political power in his hands, together with the machinery of the state, he was able to control economic institutions in the Philippines. He used the state and the economy to milk money out of the public coffers and distribute them among his cronies. Litonjua (2001) describes Marcos's crony capitalism as a "kind of subcontracting through which [Marcos's] business associates monopolized important areas of the economy for plundering by means of special taxes, production privileges, and import-export licenses." Moreover, he describes that each of the cronies had their own kingdom: "Benedicto was the sugar king, Cojuangco the coconut king, Floirendo the banana king, Campos the drug king, and, according to the buttons on the intercom system at Malacañang Palace, Ferdinand Marcos was simply the King" (Litonjua, 2001). By monopolizing each important sector in the economy, many private businessmen who are not associated with Marcos became "less willing to invest and expand in the Philippines for fear of attracting attention and instead moved their money outside the country" Dohner et al. (1989).

I argue that Marcos abused the state machinery to benefit himself and his cronies by heavily utilizing industrial policies and expansionary fiscal and monetary policies. During the beginning of his administration from 1964 to 1968, Marcos heavily focused on capital expenditure, concentrating on infrastructures.37 Although this seems a proper thing for a government to do, Marcos probably did this to make the business environment attractive and conducive to his business partners. This came at a grave price. As Litonjua (2001) notes, "the Philippine state...has substantial financial resources and broad regulatory authority which the executive branch dispenses as 'rents' to reward retainers." The government expenditure at that time rose to about 43% in real terms and its share in GDP rose from 11.5 to 14 percent (Dohner et al., 1989). However, this expenditure was heavily financed by borrowing both externally and domestically to the point when the Philippines shifted from having a budget surplus to a deficit of 3% of GDP. During 1969, government spending even increased by over 25%, and the deficit tripled in that year. Moreover, this increase in spending was financed by extensive domestic and foreign borrowing and by the central bank which increased the money supply by 20% in the last four months of 1969 alone (Dohner et al., 1989). This might be attributed to the rising inflation in 1969 to 1974, but another strong force was the oil shock of 1974. This shock costed the Philippines a real income (current account) loss of 5.6% of GDP (Dohner et al., 1989).

With the declaration of martial law, Marcos had complete control of the economy which allowed him to further raise government expenditure after the oil shock. In fact, government spending increased by 40% in 1975. This resulted in a current account deficit of over 5% of GDP. This spending, as expected, was heavily met by external borrowing (Dohner et al., 1989). Despite this increasing debt, the GDP rose as expenditure increased dramatically. Thirty percent of government spending was devoted to public investment. This was one reason why the GDP rose (Dohner et al., 1989). Since Marcos's business associates monopolized most of the important sectors of the industry, these investments benefited them greatly.

In 1980, the Philippine economy deteriorated fast after the second oil price shock. The dollar value of Philippine exports hit a peak in 1980 and fell "at an average rate of almost 5% per year through 1983" due to "weak international prices" but more importantly, "falling commodity export volumes."44 Domestic growth became slow and high global real interest rates affected major domestic firms that were mostly controlled by Marcos's cronies (Dohner et al., 1989). This resulted in a domestic financial crisis in 1981. The government chose to bailout the cronies' businesses—this favoritism, unhealthy business environment, and political climate gave rise to the Makati Business Club (MBC) in 1982 who were composed of resurgent non-crony oligarchs and businessmen in the Philippines (Hutchcroft, 1991).

The 1983 crisis marked the year of the shrinking economy and huge capital flight from the Philippines. The monopolies drove away businesses, and investors were anxious of the political climate. Debt comprised 91% of the economy (Tadem, 2018). The central bank reserves shrank by 66% which amounted to $2 billion (Dohner et al., 1989). Marcos's biggest political opponent, Sen. Benigno Aquino, Jr., was assassinated, causing an uproar among the general public and non-crony oligarchs. Despite the increasing debt and the deteriorating economy, Marcos still expanded public spending unlike his Asian neighbors like Indonesia (Overholt, 1986). He resorted to money printing, resulting in an inflation of 50.3% in the following year (Dohner et al., 1989). International creditors and the military lost faith in him, and influential members of the MBC backed his political opponents, loosening his foothold politically and economically (Hutchcroft, 1991).

Along with political and economic crises, the Philippine also experienced a moral crisis. As an illustration, Marcos's wife, Imelda, managed to flaunt her extravagant spending habit, which partly amounted to 3,000 pairs of shoes (Tantuco, 2018), even under the pressures of the economy. Resurgence against Marcos became strong, and he began losing power in the country. By February 1986, he was ousted from his position due to the EDSA Revolution.

Marcos was unable to maintain his power, because he could no longer distribute enough rents to the elite and the military. In fact, when the second oil crisis hit the economy badly, Marcos's policies worsened the situation and hurt his allies. When his regime ended, GDP began to grow as early as late 1986 as confidence in the government, private investors, and businesses slowly came back.

4 Conclusion

The culture of patrimonialism is a key element to understanding modern Filipino institutions. Shaped by the Spanish and US, the Philippines developed a patrimonial culture that deepened the system of patronage and clientelism—blurring the lines between private and official spheres (Hutchcroft, 1991).

Marcos understood and mastered this system, thus becoming the single predatory boss, or in other words, “simply, the King” (Litonjua, 2001). Most of the elite families and relatives who became his cronies received rents either in the direct form of legal monopolies or through public investments in these industries. However, Marcos financed these rents and government expenditures by printing huge sums of money and by borrowing heavily from home and abroad (Dohner et al., 1989). Together with external shocks such as the 1980 oil price crisis, the worsening political climate, and Marcos's financial mismanagement; the economy fell apart. Many passive dissenters became active political actors against Marcos, and they comprised the military, resurgent oligarchs, and even the Catholic Church. Thus, Marcos was ousted in February 1986 through the EDSA Revolution.

Extractive institutions did not start nor end with the Marcos regime as one may instinctively think. Marcos simply heightened and gravely exacerbated a deeply ingrained political culture in the Philippines—patrimonialism. The establishment of haciendas and the ilustrados paved the way for the dynastic patrimonial society that controls and extracts from the Philippines ever since it was formed in the Spanish colonial era. They are always the same players either in the foreground or background in every key historical period in the Philippines (Litonjua, 2001). As we have seen, they were also key players in Marcos's sustenance of power but also one of the reasons for his downfall. As long as dynastic patrimonialism plagues the country, extractive institutions will never disappear in the Filipino society.

References

Almeida, T. (2012). The Philippines: Oligarchipelago. *Prospect Journal of International Affairs at UCSD*. https://prospectjournal.org/2012/04/11/the-philippines-oligarchipelago/.

CuUnjieng Aboitiz, N. (2009). *Ferdinand Marcos: "Apotheosis" of the Philippine Historical Political Tradition.* University of Pennsylvania. https://doi.org/10.13140/RG.2.2.31293.46567.

Dohner, S., R., and Intal, P. (1989). Introduction to "The Marcos legacy: Economic policy and foreign debt in the Philippines.". In Sachs, J. D. and Collins, S. M., editors, *Developing Country Debt and Economic Performance, Volume 3: Country Studies - Indonesia, Korea, Philippines, Turkey*, pages 373 – 400. University of Chicago Press.

Hutchcroft, P. D. (1991). Oligarchs and cronies in the Philippine state: The politics of patrimonial plunder. *World Politics*, 43(3):414–450. https://doi.org/10.2307/2010401.

Litonjua, M. D. (2001). The state in development theory: The Philippines under Marcos. *Philippine Studies*, 49(3):368–398.

Mccoy, A. (2001). Dark legacy: Human rights under the Marcos regime. In *Legacies of the Marcos Dictatorship.*

North, D. C., Wallis, J. J., and Weingast, B. R. (2009). *Violence and Social Orders: A*

Conceptual Framework for Interpreting Recorded Human History. Cambridge University Press, Cambridge and New York.

Overholt, W. H. (1986). The rise and fall of Ferdinand Marcos. *Asian Survey*, 26(11):1137–1163. https://doi.org/10.2307/2644313.

Tadem, E. C. (2018). Philippines: The Marcos debt. http://www.cadtm.org/Philippines-The-Marcos-debt.

Tantuco, V. (2018). 3,000 pairs: The mixed legacy of Imelda Marcos' shoes. *Rappler*. https://www.rappler.com/newsbreak/in-depth/212529-imelda-marcos-shoes-mixed-legacy.

Figures

Figure 1. Philippine Inflation until 2019

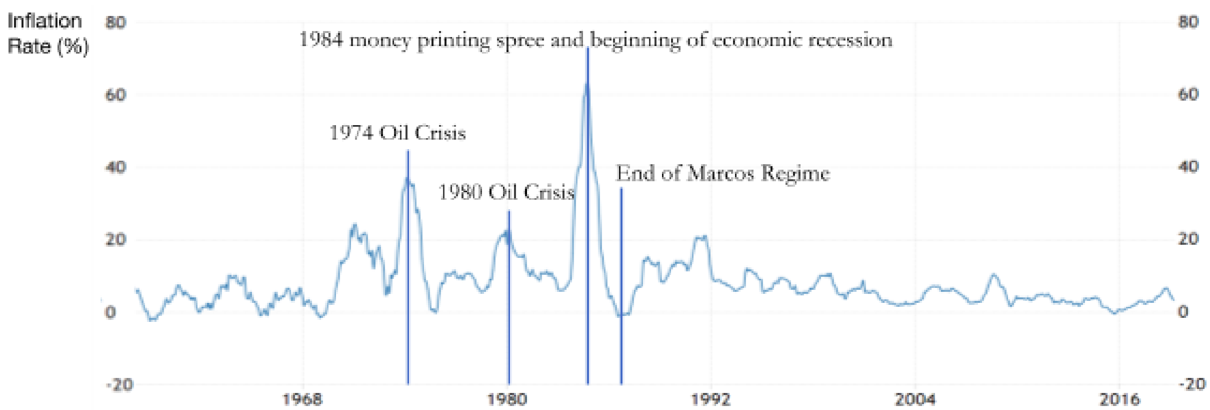
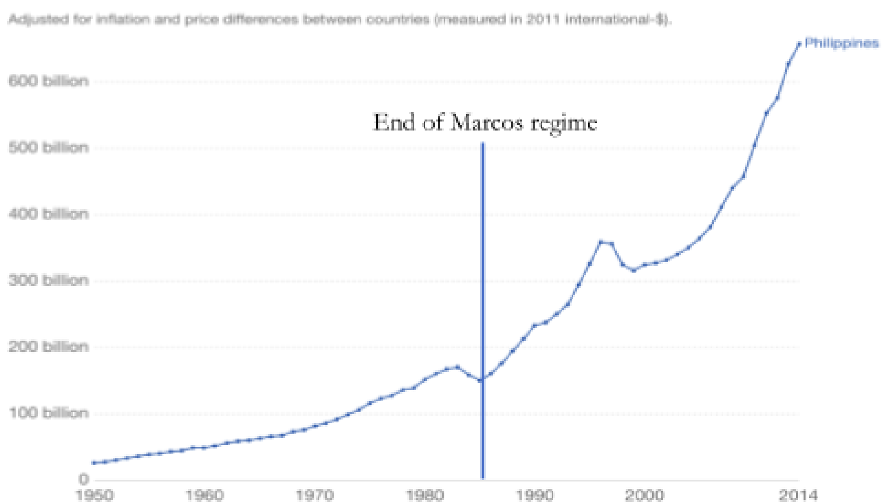


Figure 2. GDP of the Philippines



Taken from Our World in Data, "National GDP of the Philippines," Our World in Data, accessed May 7, 2019, https://ourworldindata.org/grapher/national-gdp.

Hedonic Regression Analysis for Housing Prices in the Twin Cities

Samuel Jakshtis

June 2020

Abstract

There are numerous variables that affect consumer's decisions when purchasing a home. Utilizing a hedonic regression model, the goal of this paper is to outline and determine the importance of different variables for single-family housing transactions in the Twin Cities over the course of one year. Many real estate research papers focus on broad geographical areas; however I believe that by consolidating to a micro area the results will be more accurate and specific to the Twin Cities. Moreover, the results of this paper will be comparable for other papers that are region specific in order to better understand similarities and differences in the individual region's real estate. This paper first delves into other researchers' findings and methodology, then explains the specific data used for my research. Next I explain my model, depict my results, which all lead to a discourse of the findings. This paper examines a multitude of variables and specific findings for real estate transactions of Twin City single-family homes.

Keywords: Housing; Hedonic regression; Twin Cities; Housing prices

1 Introduction

The housing industry is a dynamic and complex market that encompasses numerous variables into every transaction but is ultimately decided by individual consumer preferences and income constraints. A home purchase represents a consumer maximizing utility by prioritizing a multitude of attributes and preferences to find the optimal balance. The Twin Cities are unique for the housing industry because the area provides numerous housing choices in close proximity to one another. Many researchers have examined the value of certain characteristics and externalities that play a crucial role in determining housing prices. However, I have found no specific analyses on consumer preferences regarding single-family homes in the Twin Cities. This paper's goal is to evaluate the effects of specific housing characteristics on overall price via a hedonic regression model.

Location is a specific characteristic that is notorious for dominating real estate decisions. Different researchers have studied the effects of gentrifying a location—more green space, better transportation, and less crime—on housing prices. This research is motivated by the opportunity to study various homes' characteristics and the inherent qualities that determine subsequent prices. I believe there are other factors that are more prevalent for consumers in the Twin Cities. Utilizing a hedonic regression model, and harnessing housing data, I will attempt to build the most accurate prediction model to estimate the effects of a variety of characteristics on individual homes in the Twin Cities.

The outputs will depict significant variables when examining the overall price for home purchasers of single-family residences. This paper will first consider different researchers' methodology and findings for their specific research. Next, I will present my data and study area for my research. Following, I will introduce the regression model and the results which will lead to discourse about the implications of the outcomes. Finally, I will conclude with different caveats and directions for future research. Housing is one of the most significant investments in most people's lives and understanding the magnitude of different characteristics in homes is imperative to comprehending the dynamic patterns of consumers within the housing industry.

2 Literature Review

Any type of good or commodity can be viewed as a bundle of numerous characteristics that add or subtract to the overall value of that particular good. This is true for real estate as well; a residential property is simply a combination of characteristics (such as size, location, construction, etc.) that all contribute in some measurable way to the ultimate value that a particular buyer places on that home (Corsini, 2009). A hedonic regression analysis utilizes a function that determines the correlation and effect of each individual variable (Monson, 2009). As Rosen (1974) explains "[the] primary goal is to exhibit a generating mechanism for the observations in the competitive case and to use that structure to clarify the meaning and interpretation of estimated implicit prices" (Rosen, 1974).

Malpezzi et al. (1980) compare housing to groceries, which supports the consideration of the difficulty of quantitatively analyzing the heterogeneity in the housing markets. Sirmans et al. (2005) explain that the heterogeneity of houses means that no two houses are the same. They may be different with respect to, for example, location, size, neighborhood amenities, and quality (Sirmans, 2005). Since most of these housing characteristics are not traded explicitly, their prices cannot be observed directly on the market. The hedonic pricing model is therefore applied to estimate the marginal contribution of each property and neighborhood characteristic to the house price (Owusu-Ansah, 2011).

The intention of utilizing a hedonic regression analysis is to better gage and quantify consumer's choices in housing in the Twin Cities. Utilizing the function, one can determine the significance of each variable that can be correlated to a consumer's utility for that characteristic. As explained by Malpezzi et al (1980), "utility is not a function of commodities, but rather of the characteristics embodied in the commodities".

3 Theoretical Evidence

There are a multitude of explanatory variables in the housing market—both influential and meager—that can affect this analysis. Comparing houses in the Twin Cities is incredibly arduous due to the variation in architectural and spatial housing structures. Specifically, as explained by Pilgram and West (2018), who compared the effects of the light rail system in the Twin Cities, condos are usually associated with a central business district (CBD) area but are incredibly different from single family homes (Pilgram and West, 2018). Their study omits condominiums from the data for the purpose of singling out the effects of the light rail in the most effective manner. My study will consider condominiums because they provide information on the values consumers put on the proximity to CBDs and the certain characteristics that are preferable in a condominium. Each consumer must make their decision on housing based on their own budget constraint, demand curve, and preferences.

Considering the significance of income, Rosen (1974) explains the consumers' demand curve for housing is convex, therefore one might expect higher income consumers to purchase greater amounts of all characteristics. Only in that case, it would be true that larger income leads to an unambiguous increase in the overall "quality" consumed, and differentiated products' markets would tend to be stratified by income. However, in general, there is no compelling reason why overall quality should always increase with income. Some components may increase, and others decrease (Rosen, 1974). This study will not take individual consumer's income into account because it is assumed that the consumers will maximize utility under their specific budget constraint with the homes they purchase. The individual characteristics of each home will signify the value each consumer puts on the variables.

Goodman and Thibodeau (1998) examine the relationship between dwelling age and the market value of owner-occupied housing in Dallas, Texas. The authors theoretically proved that both depreciation and vintage effects suggest that the effect is almost certainly nonlinear and quite likely non-monotonic. Goodman and Thibodeau (1995) researched age-induced heteroskedasticity in homes and found that it is prevalent in hedonic house price equations. This finding would prove that houses may depreciate at different rates according to various conditions and attributes that may or may not be preferred by consumers. The heteroskedasticity prevalent in the age variable provides an opportunity to prove that the age of a home may follow a different function than the preconceived notion.

Consumer preferences in the housing market highlight another significant property of most heterogeneous markets: existing goods. Sheppard (1999), explains that the share of house sales accounted for by new constructions is relatively small in almost all housing markets. Consumers can substitute between them, and seek out the type, whether new or old, which maximizes utility (Sheppard, 1997). He continues to note that a technique to examine consumer choices concerning housing is to "present household choice as a function of income and the parameters of the hedonic price function have the advantage of providing (Subject to the limits of estimation) a quantitative description of actual household behavior (Sheppard, 1997).

The type of construction is impactful in housing; however, it is the individual consumer's choice of the significance of the construction. Therefore, construction type and year built can create some error in the model. Malpezzi et al. (1980) explains, "Changing construction quality can introduce another bias into the hedonic estimates of depreciation" (Malpezzi et al., 1980). This variable can be quantified by the year the house was built—age—in my research, but it is an imperfect proxy.

Controlling for the characteristics of houses is simpler than the quality, and in Pilgram and West's (2017) research they were more concerned with the homogeneity of housing through the time period for the adaption of the light rail. Therefore, homes and neighborhoods needed to be similar between time periods, and certain styles (i.e. condominiums) had to be omitted. Moreover, bias can occur when considering construction quality. I am studying a multitude of groups and their preferences of particular characteristics of single-family homes in the Twin Cities; therefore, I believe omitting different housing types would lead to bias against a certain area.

Open space is imperative in regard to my hypothesis, yet instead of directly looking at open space, it can be attributed to lot size—one of the main motivators for people to leave CBDs is for more space. As Anderson and West (2006) note "sales price rises by about 0.09% for every one percent increase in lot size, assuming the home is located an average distance from each amenity" (Anderson and West, 2006). Walking distance to amenities is an arduous concept to quantify and map with limited software. With the absence of GIS, one possibility for measuring the hedonic regression analysis is the walk score—a score that determines the walkability for occupants to complete tasks—that is associated with urban areas where fewer people rely on cars.

Proximity to amenities and walkability vary for different home buyers, therefore the importance individual consumers place on various characteristics must be considered. Reverting back to Malpezzi et al.'s (1980) metaphor of groceries, the importance of home characteristics is relative to other features of the home. The bundle of home characteristics and the importance of the individual associated characteristics will be measured via the location of the homes. For instance, as Anderson and West (2006) explain "urban residents in dense neighborhoods near the CBD place substantial value on proximity to open space, while suburban residents do not appear to value open space as highly. That is more likely because people in the outskirts of urban centers value size and space more than the urban residents who appreciate walkability to a greater degree.

The concept that land is more valuable near centers of business districts is contingent upon the demand, which, alongside the entire housing market, will increase with population. DiPasquale and Wheaton (1996) infer that higher populations in cities force people outward which makes existing developed land more valuable, while also positively affecting most land values near the CBDs (DiPasquale and Wheaton,1996). The population can be a significant determinant in a hedonic model; however, I am only using current prices with the present demand; the population will be factored into the prices that reflect the demand for the particular housing unit.

When considering home prices relative to an explanatory variable, Pilgram and West

(2017) controlled for home prices near their research area that were sold more than once, to understand the effects of the light rail system being installed (Pilgram and West, 2017). In order to describe the response variable accurately, controlling for home characteristics such as total square footage, walkability, parking, the age of the home, rooms, bathrooms, and zip code is essential.

To explain home prices, and the significance of trends in real estate in the Twin Cities, increasing the number of explanatory variables tends to—but not always—better the prediction of the hedonic regression model. However, different variables have varying importance on the overall model. This paper will be analyzing single family homes by different neighborhoods through the Twin Cities to explain consumer's relationships and preferences with central business areas.

4 Empirical Evidence

Many researchers have focused on distinct variables' effects on housing prices; however, the housing market can vary according to a number of elements and their relation to one another. Heikkila (1989) explain for their study on land values in a polycentric city, "residential property prices (in particular, the market selling prices of single-family homes) have to be used" (Heikkila et al.,1989). Market selling prices are a significant factor when determining the correct quantitative measurement for homes. Homes could be based on the assessed value (the dollar value assigned to a property to measure applicable taxes) and market value (the price paid for the house on the open market). In 2017, the Minnesota Department of Revenue (MDR) wrote in the Property Values and Assessment Report that the average home sells for 6.4% more than its assessed value (Minnesota Department of Revenue, 2018).

Locally specific negative externalities, inherent in neighborhood systems and structures, can drastically affect home prices. Crime rates in nearby neighborhoods are important factors that consumers consider when buying a home. Bowes and Ihlanfeldt (2001) reported that an additional crime per acre per year in census tracts in Atlanta decreases home prices by around 3%. Gibbons (2004) in London, found that a one-tenth standard deviation increase in the recorded density of incidents of criminal damage has a capitalized cost of just under 1% of property values. Higher crime rates tend to parallel other attributes in neighborhoods. For instance, neighborhoods with high crime also may experience fewer environmental amenities (e.g., close to parks, lakes, playgrounds, good schools), isolation (poor accessibility), proximity to major highways and transport nodes (with noise and air pollution) industrial land use or commercial/entertainment areas (e.g., close to bars, restaurants, pubs) (Ceccato and Wilhelmsson, 2011).

Alongside crimes, researchers have found that sex offenders are significant determinants for housing prices. Larsen et al. (2003) estimated that houses sell for 17% less within 0.1 miles of a sex offender, Linden and Rockoff (2008) estimated that the true change is closer to 4%. However, as Aliyu (2016) notes, the researchers did not take into account that sex offenders systematically move into cheaper houses than the average individual (Aliyu et al., 2017).

Reporting rates are known to be higher in more affluent neighborhoods (Aliyu et al., 2017); therefore, the crime rates may be biased toward higher income neighborhoods. Aliyu et al. (2017) note that crime is usually thought to be an exogenous variable, however, this is not the case; the researchers explain five scenarios where crime can be considered an endogenous variable in a regression model for housing. The first case considers lower-income neighborhoods to attract lower-income individuals, and lower income is correlated with higher crimes. Secondly, neighborhoods with more expensive homes attract criminals by offering higher expected payoffs in terms of the market value of stolen goods, and in effect, have higher crime rates. Aliyu et al.'s (2017) third point follows the second point, where crime statistics are limited to only those crimes that are reported to police.

The fourth point pertains to the aesthetics of individual homes. The researchers explain that some unobservables that increase the attractiveness of a property (e.g., large windows or a secluded backyard) also make the property an easier target for crime. Finally, self-protection for homes is attributed to higher income areas where more crime occurs and Aliyu et al. (2017) suggest that crime can be less prevalent in more affluent neighborhoods because of the deterrence provided by self-protection measures (Aliyu et al., 2017). Crime can have a significant effect on a home's value; most consumers prioritize crime and the overall well-being of the neighborhood before purchasing a home.

There are numerous ways in which researchers have analyzed characteristics of neighborhoods. Anderson and West (2006) utilize a hedonic function that estimates the price of open land on a home's value via its structure, neighborhood characteristics, and environmental amenities (Anderson and West, 2006). The authors note that the hedonic method has some imperfections when considering the bias, created from omitting unobserved variables associated with open space. They continue with a model that allows for the elasticities to determine the effect of open space on home values.

Aliyu et al. consider the aesthetics of a home in their research, and the subsequent relationship is between age—correlated with the year the home was built—and the aesthetic qualities of the home. A common notion is that new homes will sell for more because of the aesthetics and construction quality associated with newness. However, researchers have found a point when age actually increases the home's value due to its historical significance. Winson-Geideman (2011) analyzed the effect on price of a home's age in Savannah, Georgia. They explain that buyers are willing to pay a premium for the right to claim ownership of the oldest or one of the oldest homes in an area where age and history is revered, otherwise known as the "antique effect." Winson-Geideman et al. (2011) found that at 119 years old age positively affects home value. Their research concluded that—controlling for other variables—a home's depreciation schedule turns upward if the consumer views the home as an antique, thus having historical significance.

For my research, it is imperative to consider the buyers' consideration and desire to optimize utility when purchasing a home. The behavioral side of real estate economics can become an entirely new study; however, this research will assume that the response variable will be determined by the buyers—under a budget constraint—who wish to optimize utility and purchase a single-family home/condominium/townhome in the Twin Cities. The explanatory variables will determine the significance of a home's characteristics that motivate buyers in regard to the multitude of characteristics of each home.

5 Data and Study Area

An ideal data set for this study would be generated using geographical software which could map the distances of homes to a central location and utilize the distances as an explanatory variable in the regression analysis. However, I do not yet have experience using this software.

Therefore, I will use the walk-score as the primary determinant of location for individual homes. The walk score is an algorithm based on walking distances to various amenities where points are awarded for relative convenience. The walk score is based on a 0-100-point scale: a perfect score (100) is for a location that is within a five-minute (0.25 mile) radius and signals that there is no car dependency. Typically, scores of 0-30 (¿30-minute walk) have no walkability and signifies that the area is heavily car dependent. The walk score will denote the importance, or value, home buyers place on location.

Crime rates are another critical factor considered by home buyers. Researchers have noted: "Higher crime rates tend to parallel other attributes in neighborhoods" (Ceccato and Wilhelmsson, 2009). I would hypothesize as crime rate increases, home values decrease. Income per household is also a notable determinant for neighborhoods because the variable could depict the direct relationship between location, income, and housing prices. Individual allocation of income is at the core of this research and utilizing the data will show how consumers in the Twin Cities prioritize variables based on different income levels.

The construction quality of homes is one more variable that can assist when predicting the value of single homes in the Twin Cities. This variable could either be categorical or quantitative and would map out not only the desirability of older versus newer homes, but also the state of homes currently on the market. Having a standard construction quality variable for each home would allow for completeness and clarity in analyzing various single-family homes throughout the Twin Cities, as explained by Malpezzi et al. (1980).

With the exception of the walk score—which is in place for a GIS software—the data I used did not have the aforementioned specific variables, but it did have a significant amount of other characteristics that will be used in the hedonic regression analysis to predict housing prices on single-family residences in the Twin Cities. Utilizing data from NorthStar major listing service (MLS), I have gathered information from every zip code in the Twin Cities regarding the sale of single-family homes in the past year (March 11th, 2018 through March 11th, 2019). After some data wrangling, the total number of entries of single-family residences listed within that time frame is 10,284 single family houses, with 22 variables.

The first variable that ties into the regression analysis is the state of listing(s), which signifies whether a listing is active, pending, temporarily not available to show (TNAS), or sold. This will be important when filtering for houses that might have certain characteristics (i.e. why did a listing expire?). Though this variable will not be part of the regression analysis, it determines the state of the listing for the home and directs the next two variables that are crucial in this study. List date is the date on which the house became active and on the market for buyers; the off-market date occurs when the house either sells or the listing contract expires and therefore is no longer available on the open market but is kept on MLS records.

Listings will change dramatically based on certain factors, especially the seasons. In real estate, there is a strong correlation between sales and the seasons. Miller et al. (2011) find that "using a hedonic pricing model to account for housing characteristics and the standard HP-filter system to extract the trend and the cyclical/seasonality component of the prices, our findings indicate significant variation during the year for most months around the country". Understanding this phenomenon, I separate every list date into days of the year (0-365) where January, 1st is day 0 and day 365 is December 31st. Continually, I further separate the number of days in each season, as calculated by Bromberg (2016): Fall has 89 days, Winter has 89 days, Spring has 92 days, and Summer has 93 days. Working off of the varying lengths of the number of days in each season, I utilized the solstices to break the data into the appropriate seasons based on their corresponding day in a 365-day calendar. This process was not possible with the off-market date, because some houses used in the research were unsold at the time.

The municipality variable is based on zip code and is important for the externalities associated with real estate purchasing. A municipality in a city is technically a separate town, but as cities have numerous zip codes, the municipality becomes the various bureaus with different governing systems, such as school districts within a city. The municipalities are related with consumers who want to purchase homes in safe neighborhoods with good schools and effective public works. The municipality will proxy for external but unobserved factors that are arduous to quantify but are still very important for the consumer when deciding on a home to buy. Aliyu et al. (2017) explain how negative externalities such as crime and other factors will negatively affect a home's price.

The walk score is inherently tied into the various neighborhoods and, as previously stated, is the predictor for location and accessibility to CBD. This specific variable is a significant variable for this research and determines the importance that consumers assign to single-family homes' accessibility and proximity to business districts. A high coefficient will signify that being closer to the city is important, whereas a smaller, or negative coefficient will imply there are more critical factors for predicting a home's value. The walk score uses values to predict both accessibility and location, making it a broad form of analysis. It will be able to give a general-purpose prediction of a home's price relative to proximity to a CBD.

Houses have inherent physical characteristics that are core predictors for this analysis. Variables such as total finished square footage, bedrooms, and bathrooms are all fundamental considerations for buyers. One of the primary and most significant terms of the model is the style of the home the consumer is buying. Town homes attract different buyers than mid-century moderns, one buyer might love craftsman but hate colonials, or Italian stucco homes have a very particular clientele. In short, some buyers are indifferent toward style, but most consider the style of the house crucial in their purchase, which will be reflected in the style term.

Age is an imperfect, but available proxy for home condition. Utilizing a manipulated variable "age"—2019 minus the year the individual home was built—the age and age2 variables will analyze the effects of homes value as they age; moreover, the age2 will depict the individual effects of a one-year increase for homes of different ages. Every home has depreciation associated with age, and certain homes degrade quicker than others, but the age variable will provide numerical insight on the heteroskedasticity of depreciation of a home. Moreover, it will provide the possibility to analyze if there is in fact an "antique effect" on houses as they age.

The area of this study was based on the zip codes of the Twin Cities. The total count is 24 zip codes from the MLS area of Saint Paul and Minneapolis: 55102, 55103, 55104, 55105, 55106, 55107, 55108, 55109, 55114, 55116, 55117, 55404, 55405, 55406, 55407, 55408, 55409,

55410, 55411, 55412, 55413, 55417, 55418, 55419. All of these zip codes are part of the Twin Cities according to the Minnesota government. The subject area can be seen in the Saint Paul map and the Minneapolis map in the Figures section of this paper.

One important aspect about the two maps is that not all the zip codes are shown in the visual. Also, I chose to exclude some zip codes bisected by city lines (i.e. 55430) in my research, so as not to bias the results. A GIS software would allow me to gather more accurate data from the verified city lines. However, only harnessing zip codes results in a somewhat inaccurate representation of the two cities. This does not disregard the research and only justifies ways in which to better the study in the future.

6 Model and Results

Estimating the effect of these variables on home pricing using a hedonic regression analysis will provide the best possible approximation of consumers' preferences for single-family houses in the Twin Cities. The model will incorporate one interaction term that will allow for the variation in slope that is necessary to better predict the price of a single-family residence in the Twin Cities. Specifically, the style and square footage is important because consumers have different expectations of the various styles of housing options in the Twin Cities, with regards to their relative size. Moreover, the heterogeneity of housing and the subsequent size of the home will be significant because it will determine the value of each square foot consumers place for certain styles of homes. When graphing the data, it might make sense to utilize an exponential line of best fit. This would allow variation in the analysis of different variables. With price being the response variable, this study must ensure that the price is in the correct form so each variable can accurately predict the outcome.

One significant step in a regression analysis is logging the price to analyze the percentage effect on the total home price. Many researchers utilize this method; however, I will not incorporate this methodology in my primary analysis because realtors, clients, investors...etc typically explain the effects of change of a home in dollar terms, not necessarily the percentage of the total value. There are a few people that will speak in percentage terms—in terms of commonality in real estate, change is more likely to be explained in dollar terms. I will utilize the log in some instances to highlight key findings, thus the percentage and dollar terms alike will be used, but the dollar value more so. The functional form of the equation then looks like:

$$\text{Ln}(Y_h) = \beta_1 X_h + \beta_2 Y_h + \beta_3 Z_h + \cdots + \beta_n V_h + \epsilon_h$$

The model's accuracy is accounted for by the value, which is the percent of variation in the dependent variable explained by the independent variable. Through understanding the variables and the necessary actions to best predict single-family home prices in the Twin Cities, the final form of the function will take the manifestation:

$$\text{Ln}(Y_h) = \beta_S S_h + \beta_{SQ} SQ_h + \beta_{SSQ} S_h SQ_h + \beta_{BT} BT_h + \beta_{WS} WS_h$$
$$+ \beta_{SS} SS_h + \beta_{FS} FS_h + \beta_{ZC} ZC_h + \beta_A A_h + \beta_{A2} A_h^2 + \epsilon_h$$

where Y_h is the price of the individual home h; S_h is the style which is an interaction term with SQ_h, square footage; then to evaluate the significance of the interaction, the two variables act as dummy variables, thus Style and Square footage without any interaction are the next two terms in the model; BT_h is the number of bathrooms in each home; B_h is the number of bedrooms; WS_h is the walk score associated with the various neighborhoods; SS_h is the season when the home was sold; FS_h is the foundation size of the individual homes; ZC_h is the zip code that is understood as a string of characters for the regression analysis to understand the effects of one zip code relative to another—or the value of different neighborhoods; A_h is the age of the home as of 2019; A_h^2 is the age of the home squared to understand a relationship between the value of the home and its age when taking the derivative to prove that there is heteroskedasticity in home depreciation; finally, is the aforementioned error term.

Table 1 presents the results from the regression analysis on the differing effects of the variables on home prices in the Twin Cities. The three results columns show the estimates, the 95% confidence intervals, and the p-values, with the statistically significant p-values bolded. Based on the inherent qualities in a regression analysis, some values must be omitted in order to avoid multicollinearity.

When looking at seasons, the Winter variable is omitted as a baseline for Summer, Spring, and Fall. In terms of the house style, the omitted value is a 2-unit house—more commonly referred to as Duplex—that served as a basis for the rest of the styles; finally, the zip code variables omitted are both 55419 and 55102. The reason that both zip codes are omitted is due to there being too few observations in these zip codes, therefore leading to multicollinearity within the regression analysis, where the model cannot decipher between the two zip codes when comparing different effects. Omitting both is the same as omitting just one for the model.

The model returns estimates relative to these omitted categories for the different variables for housing. The means of different variables indicate the typical house sold in the sample data of houses in the Twin Cities in the last year. The mean square footage is 1675.77 ft2; the mean number of bathrooms is 1.94 per household; the mean walk score for the different neighborhoods is 65.42—which can be interpreted as having some walking accessibility, but also requiring a car; the mean foundation size is 958.99 sq. ft. The most common house sold among the numerous styles was the single-family style, with 8,625 transactions. The zip code 55106, commonly referred to as Dayton's Bluff, had the most homes sold with 866 transactions.

The intercept of the model utilizes the omitted variables to predict a 2-unit (duplex) house sold in the winter in the 55419 zip code would sell at $151,401.82. This number, however, also indicates a walk-score of 0, no bathrooms, no foundation, and 0 years old (new construction). Utilizing the model's specific figures for housing characteristics would then alter the predictions of a home's price from the intercept.

The average home price in the Twin Cities was $303,812, with the minimum price being $15,000. The maximum price a house sold for was$3,250,000. The deviation in the prices represents the various types and styles available to consumers in the Twin Cities area. Figure A is a density plot showing the prices of homes sold over the past year. The graph is strongly skewed to the right, which makes sense because residential real estate becomes more infrequent as homes get more expensive in most markets.

One of the most interesting outputs is the difference in prices throughout zip codes in the Twin Cities. Figure B represents the average home prices in the Twin Cities across the many zip codes. The most expensive zip code is 55410 and the cheapest is 55411. The disparity of the zip codes is indicative of not only the many homes' characteristics, but also the aforementioned externalities associated with each neighborhood.

A surprising find is the average price of a home sold throughout the four seasons. The disparity in average home price is represented in figure C. There are a few reasons explaining the differences in prices between the seasons. One is that it could be a representation of a greater trend of average home prices falling throughout the year. Another explanation could be in the houses themselves: if a seller knows that their house is desirable and in good shape, they will not consider the season as significantly as people who choose to wait until the more agreeable seasons to maximize their curb appeal to potential buyers. This might explain the counterintuitive finding that winter prices are higher on average than other seasons' prices.

The model yields many notable estimates. Firstly, the single-family-stand-alone unit, with a townhouse-type homeowner association, sold for the most compared to the other styles. The zip code with the most positive impact on the intercept was the 55405 neighborhood, but homes in that area did not sell for the most on average, meaning residences in that zip code did not have the other desirable qualities for homes in the Twin Cities.

One of the most interesting estimates is the effect of the walk score, which is negative, or has an adverse effect on home values. This shows that consumers of real estate in the Twin Cities focus on qualities found outside of central business districts. Moreover, this is notable because many would assume a higher walk score would lead to more expensive homes. However, in the Twin Cities, consumers value other variables more than walk score. These other variables are found in neighborhoods that are farther from central business districts. Walk score may also proxy for negative neighborhood attributes I do not include here, such as noise or crime rates, which are higher in dense areas with greater walk scores.

Another interesting result was the relationship between price and age. Age and age squared are two significant variables in this research and are designed to map the price of a home versus its age as of 2019. Utilizing data from the past year (March 2018–March 2019), I plotted an average home's value as it aged to analyze the heteroskedasticity of depreciation. I classify an average home by the mean of quantitative variables, or the most entries within a categorical variable, and add their values to the original intercept.

Single-family was the most sold style of home, with 8,625 transactions, which adds \$112,084 to the intercept. The average square footage is 1,675.774 ft2 per home sold, and every additional square foot adds \$220.06. The interaction between style and square footage is -73.86 sq. ft. , which is multiplied by the average square footage of 1,675.774 sq. ft. The average number of bathrooms per house is 1.94, and every additional bathroom adds \$25,929.41. The average walk score for homes in the Twin Cities was 65.424, and every increase in walk score decreased the home's value by \$2,124.28. The most homes sold were in Spring and the particular season decreases the price by \$7,335.74. The average foundation size per home in the Twin Cities was 958.998, and every square foot in foundation size added \$33.61 to the total value of the home. Finally, the most transactions occurred in the 55106 zip code, which decreased the average home value by \$192,503.09.

Utilizing the equation to find the mean home in the Twin Cities with an age of zero years resulted in an average house with a price of about \$223,329.48. Continually, the equation to graph the relationship between price and age is the entire regression model, including the age and age-squared variables:

$$\text{Price}_h = \$223,329.48 - (\$1,677.81 \cdot \text{age}_h) + (\$9.18 \cdot \text{age}_h^2)$$

The resulting graph is shown in Figure D.

Utilizing the derivative of the function, I found that 91.38 years is the point at which a home's value is minimized by its age. This is significant because it explains the point at which a home's value stops decreasing in the eyes of consumers and begins to appreciate because of its historical qualities. These numbers are not adjusted for inflation because they represent homes sold in the last year. The homes' trends are based on the year they were built and the price for which the individual homes sold. It is crucial to note that these findings are for a typical home in the Twin Cities. Moreover, age is not interacted with any terms, and the age-price relationship could differ across home styles—among other categories—in a way for which this regression does not account.

The lower price reflects that consumers depreciate a home's worth based on the year it was built up to 91 years; at that point the historical factor begins to appreciate the home's value, because the consumer sees the value in owning an older home. 91 years is significantly less than the 119 years found in Winson-Geideman et al.'s (2011) research. There are a few reasons that could explain this discrepancy: 1) they studied only historical districts in Savannah, Georgia, therefore their study area was biased toward older homes, and the age of a home would inherently be older than analyzing against newer homes; 2) the actual historical value of a home is subjective to the buyer, and different buyers in separate parts of the countries place varying values on the historical significance of homes. My data from the Twin Cities over the past year shows that buyer preferences start to favor the historical value of a home after 91 years, reversing its depreciating schedule.

The regression model's R^2 is 0.755 and the adjusted R^2 is 0.754, meaning the model and independent variables explain about 75% of the variation in the dependent variables. Although the model explains a significant portion of the variation, the remaining 25% is indicative of consumer preferences that are incredibly arduous to explain in a quantitative method. Many real estate transactions are intertwined with emotions, but regardless, the model does a considerable job at depicting the major considerations for real estate consumers in the Twin Cities. The next section provides an overview and a discourse for further research.

7 Conclusion

Home prices in the Twin Cities are determined by a plethora of consumer preferences reflected in home's price. The hedonic regression model highlights significant variables that consumers value when they purchase a single-family home. Different zip codes, walk scores, and various styles of homes represent some of the variables that are considered in the purchase. This paper explored other researchers' methodology and results from similar studies. Following, I explained the data utilized alongside the specific model for predicting home prices in the Twin Cities. Finally, I found specific findings related to zip codes, seasons, and the age of individual homes.

One of the main dilemmas with this research is the time frame analyzed. Though there were a significant amount of transactions, one year only provides a snapshot to the entirety of the housing market. For instance, winter had the highest average sale price among the seasons. This is a significant finding, but with only one of each season, the results cannot be as representative nor accurate of the population of houses sold over time. Moreover, utilizing more data over a longer time frame would provide insight to trends across a period of years and visualize the dynamic preferences of consumer biases of single-family homes.

This research was also subject to a small geographic concentration of only the Twin Cities. Although there are advantages to examining a micro area, there can be some cons. The trends found in this research are not representative of the state of Minnesota nor the United States. Housing markets change frequently, and there is little to no ubiquity when generalizing all of the United States in terms of real estate. Continually, the Twin Cities are an urban center, analyzing rural areas creates the possibility of seeing varying trends and estimates. Also, including rural areas could lead to an over-generalization of predictions, thus one would have to separate urban and rural areas if conducting a similar study.

The most important part of continuing this research would be based around GIS software, which would allow one to gage distance of individual houses from CBD. The distance could then be used to better compare the prices of homes sold based on their geographic location. There are also numerous externalities within geographic locations. Aforementioned in the data section of this research paper: crime rates and income. Higher crime rates will hypothetically decrease the price of homes, however, utilizing the GIS would showcase different interactions that could potentially offset the higher crime rates. Separating consumers based on their incomes would lead to multiple opportunities because one could research the shifting trends based on income levels or, more specifically, budget constraints.

This research found significant trends that are important to consider regardless of the size and time frame of the sample size. The price of a home relative to its age is heteroskedastic; I found that at 91 years of age, consumers see value in the historical aspect of the home. This finding is based on the average of all variables in the regression model, and it should be noted that there could be specific changes that would skew this number. Also, there are externalities in the number that are not considered when analyzing the age of a home and its relative price.

There are numerous other factors to consider when looking at consumer preferences of single-family homes in the Twin Cities. One significant finding is that increasing the walk score—or closeness to a CBD—decreases the relative price of a single-family home. This finding is significant because the negative association with price proves consumers value attributes associated with homes that are distant from central business districts. There may be confounding variables in this finding, but the overall idea is important when considering the value home buyers place on proximity to a CBD. Consumers of single-family homes in the Twin Cities place a higher value on characteristics found in neighborhoods further from central business districts, ergo the relationship between home price and relative distance to CBD is negative.

Utilizing the hedonic regression method to understand various components of the Twin Cities housing market provided valuable quantitative insight. There are many ways to further this research, not only in the Twin Cities, but anywhere with real estate transactions. Mapping the qualitative measures of buying a home is close to impossible, however, the emotions of buying a home are significant factors. Regardless, the price of a home in the Twin Cities in terms of the inherent qualities the house possesses are well defined through this hedonic regression model.

References

Aliyu, A. A., e. a. (2016). Impact of crime on property values: Literature survey and research gap identification. *African Scholar Journal of Humanities and Social Science*, pages 68–99.

Anderson, S. T. and West, S. E. (2006). Open space, residential property values, and spatial context. *Regional Science and Urban Economics*, 36(6):773–789.

Bowes, D. R. and Ihlanfeldt, K. R. (2001). Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics*, 50(2):1–25.

Bromberg, I. (2016). The lengths of the seasons (on Earth). http://individual.utoronto.ca/kalendis/seasons.htm.

Ceccato, V. and Wilhelmsson, M. (2011). The impact of crime on apartment prices: Evidence from Stockholm, Sweden. *Geografiska Annaler. Series B, Human Geography*, 93(1):81–103.

Corsini, K. R. (2009). Statistical analysis of residential housing prices in and up and down real estate market: A general framework and study of Cobb County, GA. Master's thesis, Georgia Institute of Technology.

DiPasquale, D. and Wheaton, W. C. (1996). *Urban Economics and Real Estate Markets*. Prentice-Hall, Inc. 35–59.

Gibbons, S. (2004). The costs of urban property crime. *The Economic Journal*, 114(499):F441–F463.

Goodman, A. C. and Thibodeau, T. G. (1998). Dwelling age heteroskedasticity in repeat sales house price equations. *Real Estate Economics*, 26(1):151–171.

Heikkila, E. e. a. (1989). What happened to the CBD-distance gradient? land value in a polycentric city. *Environment and Planning A*, 21(2):221–232.

Linden, L. and Rockoff, J. E. (2008). Estimates of the impact of crime risk on property values from Megan's Laws. *American Economic Review, vol. 98, no. 3, June 2008*, 98(3):1103–1127.

Malpezzi, S., Ozanne, L., and Thibodeau, T. (1980). *Characteristics of Prices in Fifty-Nine Metropolitan Areas*. Department of Housing and Urban Development Office of Policy

Development and Research. https://www.huduser.gov/portal//Publications/pdf/HUD-50814.pdf.

Minnesota Department of Revenue (2018). 2018 property values and assessment practices report – assessment year 2017.

Monson, M. (2009). Valuation using hedonic pricing models. *Cornell Real Estate Review*, 7:62–73.

Owusu-Ansah, A. (2011). A review of hedonic pricing models in housing research. *Journal of International Real Estate and Construction Studies*, 1(1):19–38.

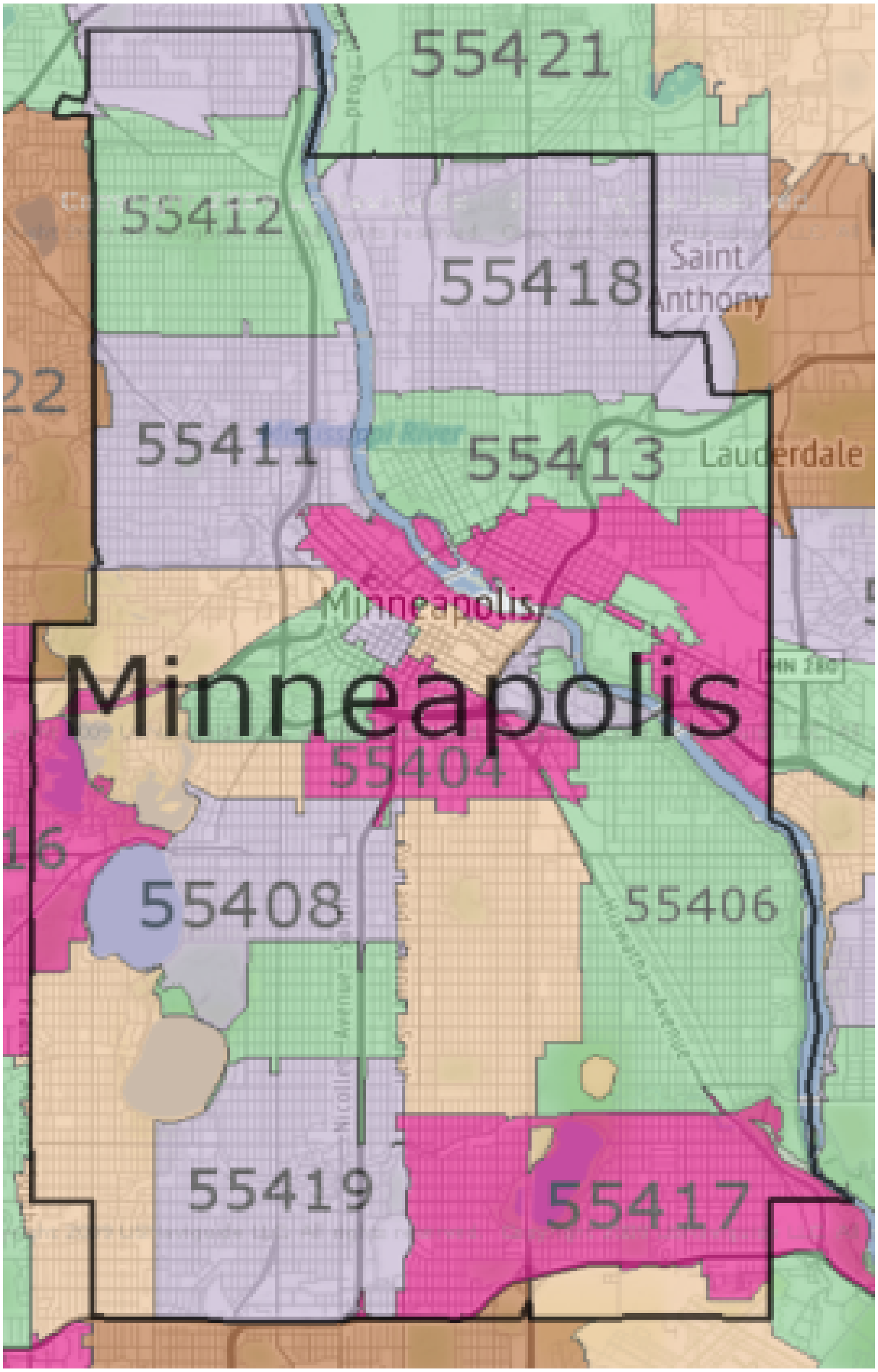
Pilgram, C. A. and West, S. E. (2018). Fading premiums: The effect of light rail on residential property values in Minneapolis, Minnesota. *Regional Science and Urban Economics*, 69(C):1–10.

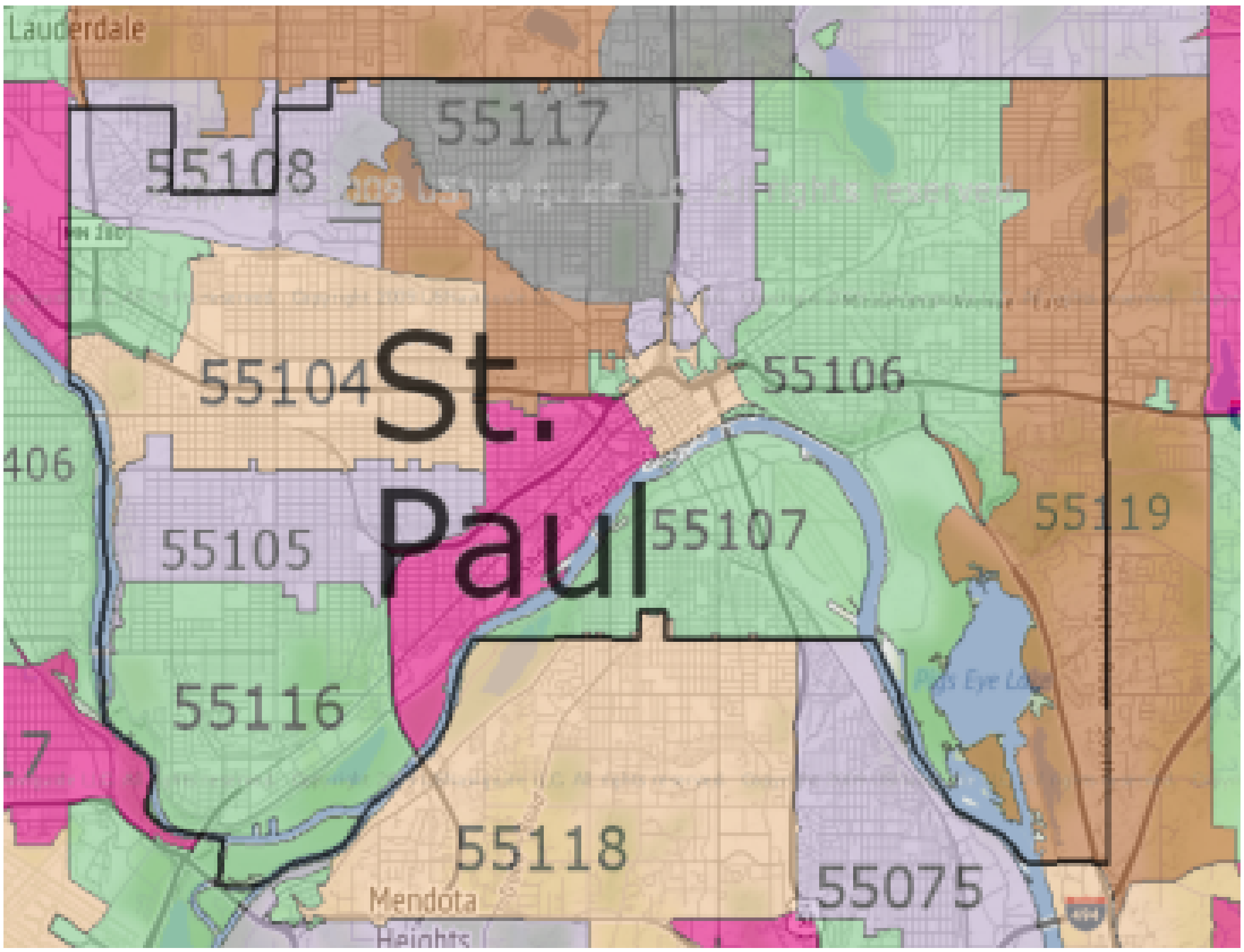
Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *The Journal of Political Economy*, 82(1):34–55.

Sheppard, S. (1999). Hedonic analysis of housing markets. In Cheshire, P. and Mills, E., editors, *Handbook of Regional and Urban Economics, Vol. 3*, page 1595–1635. Elsevier.

Winson-Geideman, K. e. a. (2011). The impact of age on the value of historic homes in a nationally recognized historic district. *Journal of Real Estate Research*, 33(1):25–48.

Figures


55421
55412
55418
Saint
Anthony
22
55411
Mississippi River
55413
Lauderdale
Minneapolis
Minneapolis
55404
16
55408
55406
Nicollet Avenue
Hiawatha Avenue
55419
55417


Lauderdale
55117
55108
55104
St.
Paul
55106
406
55105
55107
55119
55116
55118
Mendota
Heights
55075

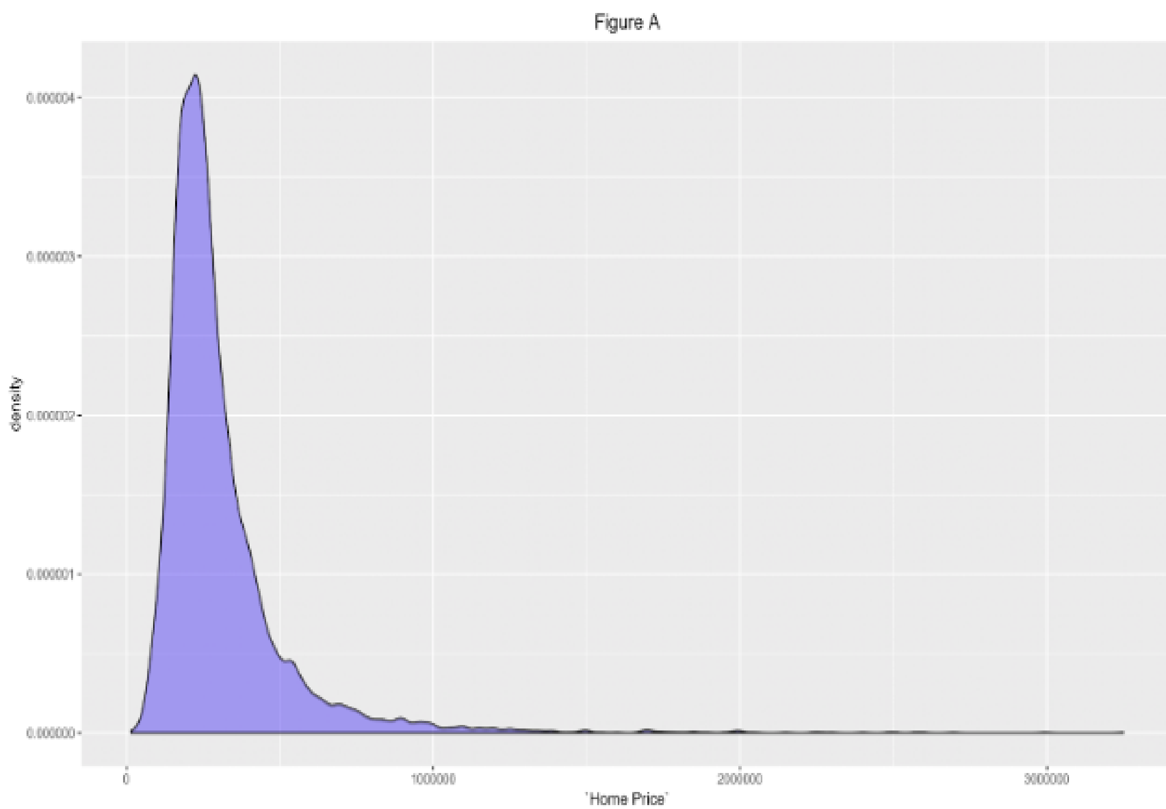


Figure A
0.000004
0.000003
0.000002
0.000001
0.000000
density
0
1000000
2000000
3000000
`Home Price`

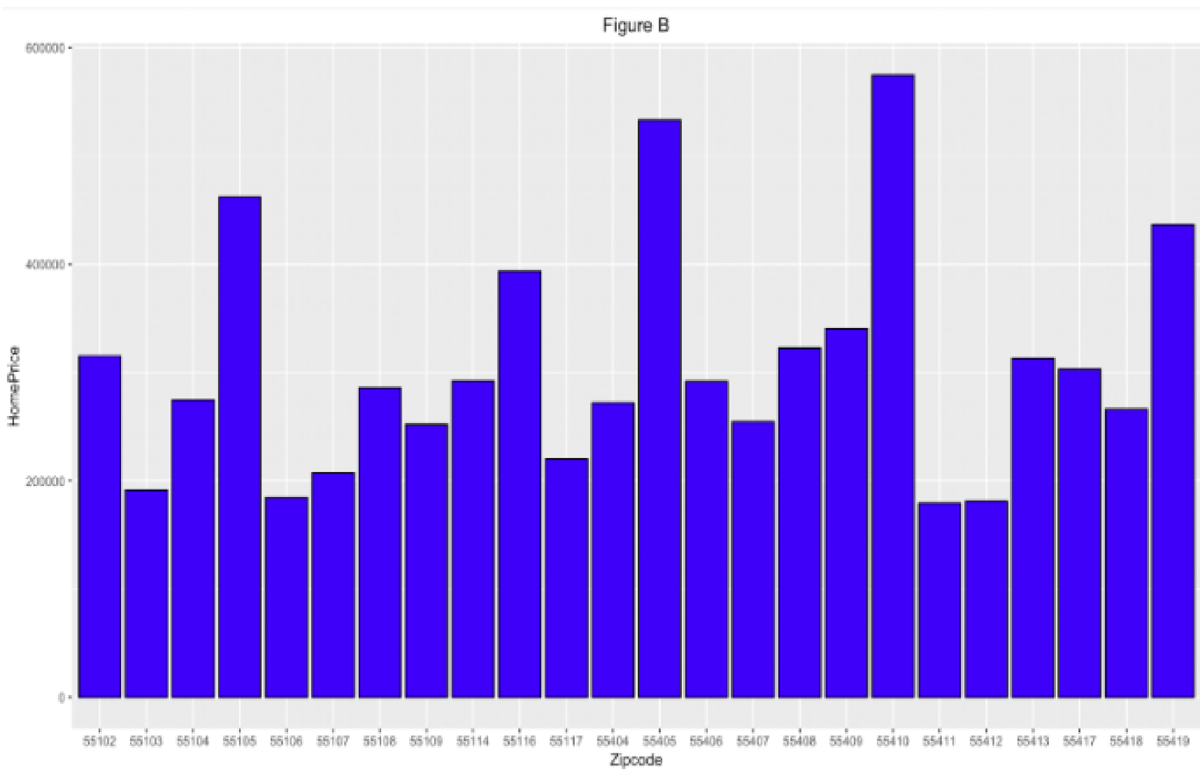


Figure B
600000
400000
200000
0
HomePrice
55102
55103
55104
55105
55106
55107
55108
55109
55114
55116
55117
55404
55405
55406
55407
55408
55409
55410
55411
55412
55413
55417
55418
55419
Zipcode

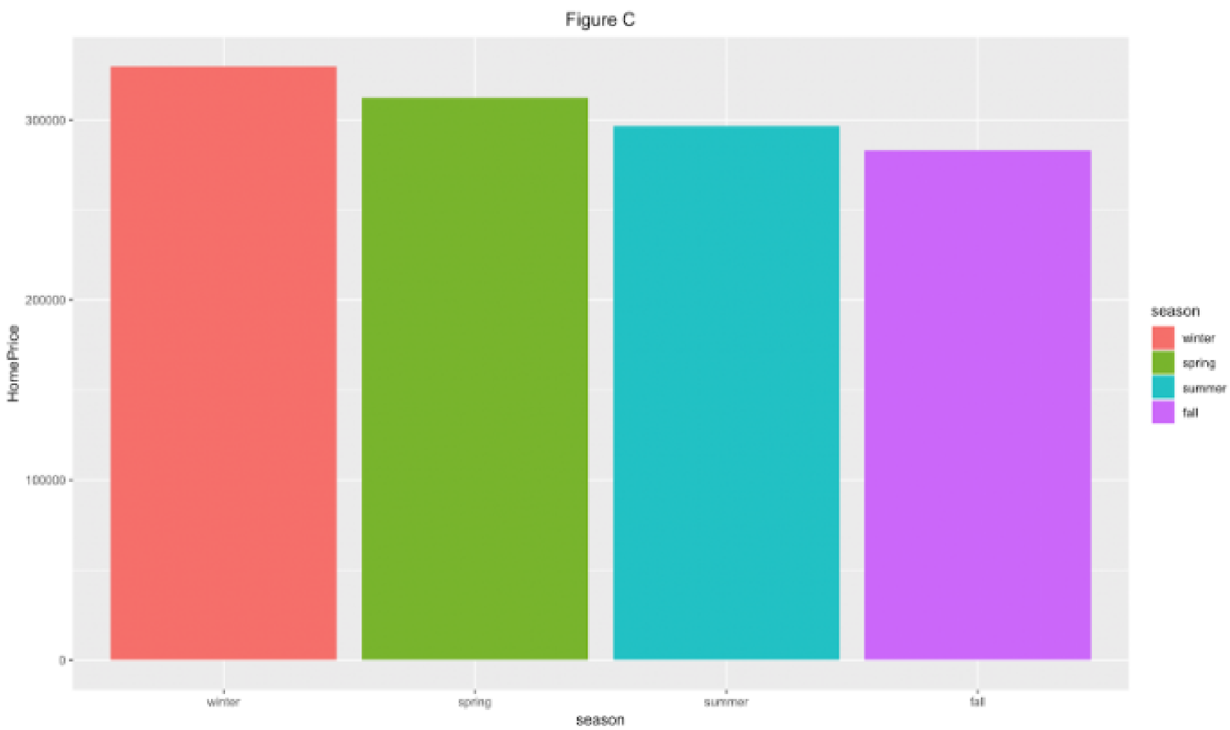


Figure C
300000
200000
100000
0
HomePrice
winter
spring
summer
fall
season
season
winter
spring
summer
fall

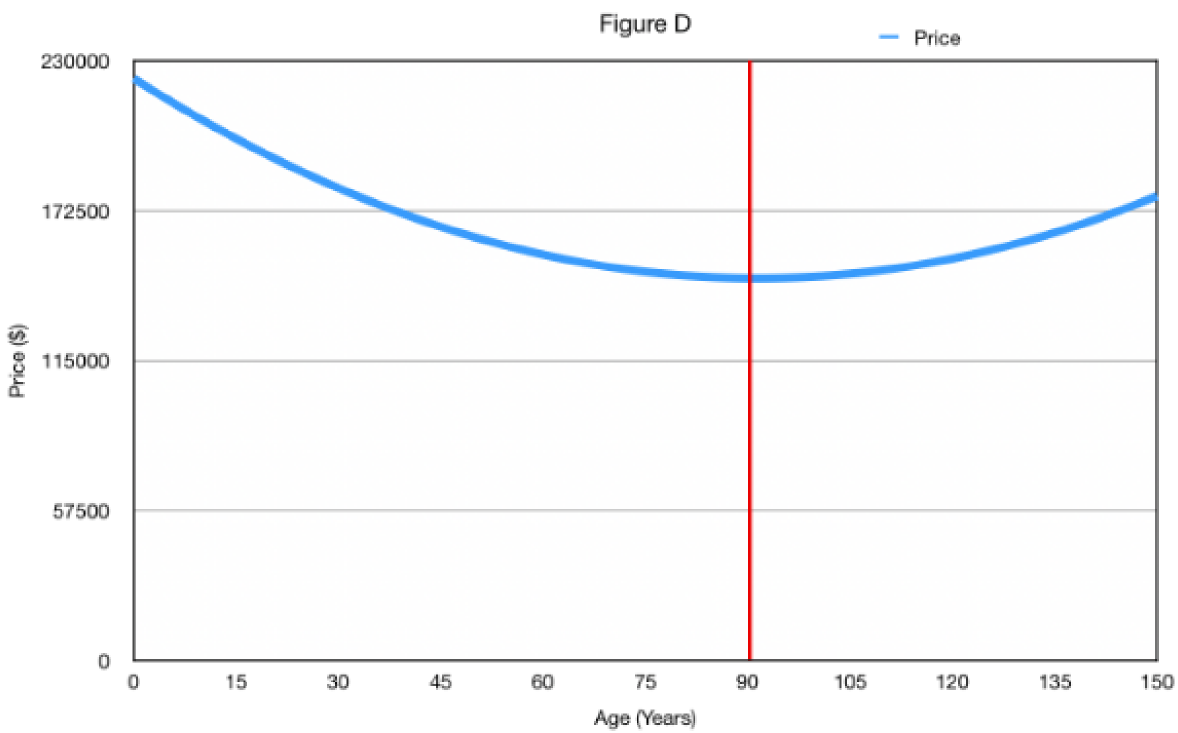


Figure D
Price
230000
172500
115000
57500
0
Price ($)
0
15
30
45
60
75
90
105
120
135
150
Age (Years)

Tables

Table 1

	Home Price		
Predictors	Estimates	Confidence Interval (95%)	P-Value
(Intercept)	151401.82	-76561.06 – 379364.71	0.193
Total Finished Square Footage	220.06	102.90 – 337.21	**<0.001**
Style-CONVERTED MANSION	125961.56	-108063.13 – 359986.24	0.291
Style-HIGHRISE	-44327.69	-270611.81 – 181956.43	0.701
Style-LOWRISE	19373.96	-206811.82 – 245559.75	0.867
Style-MANOR/VILLAGE	156012.69	-160216.77 – 472242.15	0.334
Style-QUAD/4-CORNERS	-4721.85	-368941.10 – 359497.40	0.980
Style-SINGLE FAMILY	112084.60	-113051.53 – 337220.74	0.329
Style-SINGLE FAMILY (with townhouse-type homeowner association)	218622.46	-116169.72 – 553414.63	0.201
Style-TWINHOME	173048.81	-81916.53 – 428014.16	0.183
Style-SIDE BY SIDE	-16159.00	-244280.79 – 211962.78	0.890
Bath - Total	25929.41	22126.15 – 29732.67	**<0.001**
Walk score	-2124.28	-2581.69 – -1666.88	**<0.001**
Spring	-7335.74	-13394.78 – -1276.70	**0.018**
Summer	-17206.45	-23444.22 – -10968.68	**<0.001**
Fall	-18603.46	-25098.32 – -12108.61	**<0.001**
Foundation Size	33.61	24.31 – 42.90	**<0.001**
Zipcode55103	-172097.08	-194749.56 – -149444.60	**<0.001**
Zipcode55104	-107999.48	-120662.41 – -95336.56	**<0.001**
Zipcode55105	-2940.33	-14903.89 – 9023.22	0.630
Zipcode55106	-192503.09	-206982.39 – -178023.80	**<0.001**
Zipcode55107	-147323.88	-163948.45 – -130699.31	**<0.001**
Zipcode55108	-64679.87	-80380.80 – -48978.93	**<0.001**

Table 1 (cont'd)

Zipcode55109	-119199.97	-132514.08 – -105885.86	**<0.001**
Zipcode55114	-70150.39	-98721.52 – -41579.26	**<0.001**
Zipcode55116	-73690.96	-89424.48 – -57957.43	**<0.001**
Zipcode55117	-150411.81	-161959.62 – -138864.00	**<0.001**
Zipcode55404	9615.25	-7075.80 – 26306.30	0.259
Zipcode55405	126200.75	107022.58 – 145378.92	**<0.001**
Zipcode55406	-6859.66	-17640.93 – 3921.61	0.212
Zipcode55407	-66251.55	-77121.90 – -55381.21	**<0.001**
Zipcode55408	7961.55	-4680.53 – 20603.62	0.217
Zipcode55409	-2742.51	-17279.15 – 11794.13	0.712
Zipcode55410	93751.10	82932.08 – 104570.12	**<0.001**
Zipcode55411	-174603.05	-186851.06 – -162355.03	**<0.001**
Zipcode55412	-130575.23	-141631.53 – -119518.93	**<0.001**
Zipcode55413	1694.57	-14434.49 – 17823.63	0.837
Zipcode55417	-51642.42	-62514.75 – -40770.08	**<0.001**
Zipcode55418	-143290.46	-161281.61 – -125299.32	**<0.001**
AGE	-1677.81	-1954.37 – -1401.25	**<0.001**
AGE^2	9.18	7.30 – 11.06	**<0.001**
Style-CONVERTED MANSION:Total Finished Square Footage	-84.38	-205.21 – 36.44	0.171
Style-HIGHRISE:Total Finished Square Footage	79.84	-38.80 – 198.48	0.187
Style-LOWRISE:Total Finished Square Footage	-8.60	-127.25 – 110.06	0.887
Style-MANOR/VILLAGE:Total Finished Square Footage	-141.96	-389.97 – 106.05	0.262
Style-QUAD/4-CORNERS:Total Finished Square Footage	-64.99	-303.91 – 173.93	0.594
Style-SINGLE FAMILY:Total Finished Square Footage	-73.86	-190.86 – 43.13	0.216

Table 1 (cont'd)

Style-SINGLE FAMILY (with townhouse-type homeowner association):Total Finished Square Footage	-169.86	-336.13 – -3.59	**0.045**
Style-TWINHOME:Total Finished Square Footage	-145.26	-277.68 – -12.83	**0.032**
Style-SIDE BY SIDE:Total Finished Square Footage	-21.79	-140.58 – 97.00	0.719
Observations	10284		
R^2 / Adjusted R^2	0.755 / 0.754		

Identifying and Measuring Transshipment in the U.S.-China Trade War

Yannick Laurent

December 14, 2019

Abstract

The trade war between the United States and China has caused significant economic losses to both parties. The literature indicates that China has used transshipment to avoid tariffs in the past so it is likely that such activities are occurring once more. In this paper, I seek to identify and measure transshipment in the context of the trade war between the United States and China. I find statistically significant evidence that China is rerouting almost $800 million in exports through a pool of eight intermediary countries so as to avoid the American tariffs that are levied on them. These results represent a baseline that can be used to inform future works on the issue of transshipment in the U.S.-China trade war.

1 Introduction

In March of 2018, the United States imposed 25% and 10% tariffs on Chinese steel and aluminum respectively. This marked the beginning of a tit-for-tat trade war that has raged on for the past year and a half with both sides incurring significant economic loses. In 2018 alone, American consumers experienced a total deadweight loss in excess of $6.9 billion; importers lost $12.3 billion in revenue via transfers to the government; and Chinese imports decreased by $165 billion (Amiti et al., 2019). We therefore expect that exporters will attempt to avoid these escalating costs by any means possible. This can be done in one of two ways depending on how firms choose to export. If FDI is used, one would expect actors to relocate production to a similar country that has lower trade barriers.

When direct exporting is used, firms may elect to redirect trade through intermediary countries which are subject to lower tariffs (Fishman et al., 2008). Such practices, referred to here as transshipment, are typically characterized by the shipping of goods to an intermediary country where they are, at most, insignificantly altered before being re-exported to their final destination as exports from the intermediate country. For this paper, the term "insignificantly altered" refers to changes which do not affect a given good's HS 6-digit classification. Tariffs can be avoided in this way as the final importer will apply the duty associated with the intermediary country instead of the one that is associated with the true country or origin. Transshipment thus hinges on the differences in tariff levels between importers.

This paper explores the effects of the U.S.-China trade war on transshipment to determine its presence and significance. There are two important implications of tariffs being avoided through trade redirection. First, transshipment would call into question the efficacy of tariffs as a foreign policy tool since they can be mitigated. Second, the redirection of trade could imply that the fears of an impending global economic downturn due to the trade war may be overblown. Although both of these issues have been addressed by the existing literature, there is a lack of academic research specifically analyzing transshipment in the context of the ongoing events.

Transshipment has been used as a tool for avoiding tariffs since at least the 1990s. Fishman et al. (2008) found that Hong Kong has traditionally served an important role in facilitating indirect trade to China with the express purpose of avoiding tariffs. Furthermore, the authors found that products were more likely to be transshipped when tariffs on those goods increased. Goods that were not subject to this tariff increase did not undergo any change in their transshipment levels (Fishman et al., 2008). This behavior is prevalent despite the fact that it is considered illegal as goods are supposed to be labeled according to their country of origin and not their country of last passage.

Furthermore, Fung (1998) found that, in 1993, 67% of Chinese exports and 34% of imports were re-directed through Hong Kong. This result meant that the U.S.-China trade deficit (measured using U.S. data) was off by 35% compared to reality. This highlights not only the fickleness of measuring transshipment, but also the fact that such activities can cause significant discrepancies in official data. Fung also found that the majority of imports from Taiwan to China occurred as re-exports through Hong Kong (Fung, 1996). It is worth noting that these results are specific to their time as the period in question saw China's entrance into the WTO as well as the handover of Hong Kong in 1997. While the 67% and 34% statistics are unlikely to reflect current transshipment levels, they still demonstrate the prevalence of this activity as well as the means by which it has historically been carried out.

The redirection of trade through intermediaries to avoid trade costs is only one half of the overall story; the question still remains as to which destinations one should use as middle men. This decision hinges on a variety of factors including relative tariff levels and geographical positioning. Basic intuition suggests that lower trade barriers and closer physical proximity would make a destination more attractive. Furthermore, a paper by Kokko et al. (2014) found that the presence of weak institutions hindered trade and shortened the duration of trade relationships. This study was conducted in the context of a shift in trade patterns towards non-OECD countries which was caused by these economies' relatively rapid recoveries from the 2008 financial crisis.

The authors also found that firms that were either small or belonging to industries that rely heavily on good business relationships were especially vulnerable. Based on these conclusions, trade should be redirected through countries that are geographically close to either the origin or destination, benefit from favorable trade relations with both of these countries, have strong institutions, and have good preexisting relationships with the exporter.

Another potential factor in determining which countries serve as intermediaries may lie in China's current push to fortify trade relations with its neighbors. Jacks and Novy's study of interwar trade patterns found that the erection of trade barriers during this time period mainly resulted in the fortification of existing trade blocks. More importantly, the authors drew a parallel between these findings and current events to suggest that the U.S.-China trade war may solidify existing and emerging trade blocks (Jacks and Novy, 2019). This fits in nicely with the Chinese government's current push for bilateral trade agreements with its neighbors, making these countries likely targets for transshipment.

I find statistically significant evidence in support of the hypothesis that transshipment is occurring in the U.S.-China trade war. Measurement of this activity indicates that almost $800 million worth of Chinese exports to the United States are being rerouted through eight intermediary Asian countries. This represents approximately 0.15% of total American imports from China in 2018. The remainder of the paper is organized as follows. Section II provides a brief overview of the trade war, Section III details the empirical approach that I use to measure transshipment, Section IV describes the data, results are discussed in Section V, and Section VI concludes.

2 Overview of the Trade War

In order to properly discuss transshipment in the context of the trade war, it is important to first understand its underlying facts and motivations. According to President Trump, his administration's decision to impose tariffs was caused by a desire to end China's supposedly unfair trade practices, reduce the United States' trade deficit, and promote national security (Tan and Chen, 2019). Therefore, the trade war appears to be largely intended as a pressure tactic to force Beijing to implement domestic reforms related to these issues.

Unlike the Unites States, China's economy is heavily controlled by the central government which establishes five-year plans that aim to achieve a series of targets, often emphasizing economic expansion (Ishimine, 1979). The government's current plan continues this legacy through the Belt and Road Initiative, the establishment bilateral trade agreements in the Asia-Pacific region, and support for the domestic manufacturing, telecommunications, and information technology industries (Tan and Chen, 2019).

A year and a half of tariff increases between the world's two largest economies has resulted in losses for both parties. To put this into perspective, 2018 exports from China to the United States were on the order of half a trillion dollars, 237 billion of which were subject to tariffs by the end of that year (Tan and Chen, 2019). Even though estimates point to China losing most in the short run, the medium-term impact to the United States will eventually close the gap. The net result is that both economies are expected to decline in size by 0.5% each over the next ten years (Tan and Chen, 2019). This being said, the current behavior of each government indicates that they are willing to put up with these losses as long as doing so allows them to avoid making large concessions to the opposition.

A second, more comprehensive, analysis of the issue found that there are a few winners in the trade war. Whereas American and Chinese GDP are expected to fall by 0.2% and 1% respectively over 2-3 years (in the worst-case scenario that sees all threats and retaliations carried out), developing Asian countries can expect modest benefits from trade redirection towards them. However, developed regions such as Japan and Europe will suffer heavily as they will be disproportionately impacted by tariffs on autos and auto parts (Abaid, 2018).

3 Model

The main issue that arises when measuring transshipment is the near impossibility of tracking individual containers from their port of origin to their final destination. This is problematic for several reasons. First, there is a distinct possibility that Chinese exporters are illegally relabeling their goods as having originated from a third country while still shipping them out of China for the purpose of getting around American tariffs (Fishman and Wei, 2004). Such activity would only be measurable by physically going to Chinese ports and inspecting each individual shipment. Therefore, the degree and significance of illegal relabeling is virtually impossible to measure given the resources available.

A second major measurement hurdle is the likelihood that American importers are choosing to source a higher portion of their imports from third countries and that Chinese producers are choosing to shift exports to a third country as a result of the tariffs. This would muddy the data as an observed increase in American imports from the intermediaries could represent any combination of transshipment and legitimate exports from those countries. The same issue is present when analyzing differences in trade between these countries and China.

This paper will rely on a time-series regression technique to estimate transshipment of Chinese goods to the United States. Each regression will take on the form:

$$Trade_{i,j,t} = \beta_1 Tariff_{i,j,t} + \beta_2 TSavings_{i,j,t} + \beta_3 G_{i,t} + \beta_4 C_{i,t} + \alpha_i + \alpha_t + \epsilon_{i,j,t} \tag{1}$$

where i denotes a third country other than the United States or China, j denotes different goods, and t denotes time. $TSavings_{i,j,t}$ represents the tariff savings that are incurred by engaging in transshipment rather than direct exports. This is derived by adding the tariff imposed on China by country i on good j to the tariff imposed by the United States on imports of good j from country i. This sum is then subtracted from the American tariff on good j from China. The rapid escalation and evolution of the trade war means that tariff savings vary with time. In addition to being the main measure of the motivation to engage in transshipment, tariff savings also indirectly capture the existence of possible bilateral trade agreements between China and intermediary i. This is important since Jacks and Novy (2019) suggest that trade barriers enhance trade between partners that have preexisting trade agreements.

The variable $G_{i,t}$ stands for the quality of institutions in country i, proxied using the Ease of Doing Business Index. This variable's beta is expected to have a positive sign which would be consistent with the findings of Kokko et al. (2014). $C_{i,t}$ represents country i's capacity to handle the large increase in trade activity that would result from the occurrence of transshipment. This is proxied using the perceived quality of port infrastructure. ? measures industry fixed effects and ? is a time fixed effect. ? is an error term. Individual trade capacities and institutions are initially used instead of country fixed effects out of a desire to gain some insights as to which factors contribute most heavily to a country being used as an intermediary.

In order to circumvent the issue caused by the sourcing of goods from country i as a result of tariffs, this paper will rely on three separate regressions. Each of these will use the functional form outlined in equation 1 and are structured as follows:

$$Trade_{ch,us,j,t} = \beta_1 Tariff_{us,ch,j,t} + \beta_2 G_{us,t} + \beta_3 C_{us,t} + \alpha_j + \alpha_t + \epsilon \tag{2}$$

$$Trade_{i,us,j,t} = \beta_1 TSavings_{i,j,t} + \beta_2 Tariff_{us,i,j,t} + \beta_3 G_{i,t} + \beta_4 C_{i,t} + \alpha_j + \alpha_t + \epsilon \tag{3}$$

$$Trade_{ch,i,j,t} = \beta_1 TSavings_{i,j,t} + \beta_2 Tariff_{i,ch,j,t} + \beta_3 G_{i,t} + \beta_4 C_{i,t} + \alpha_j + \alpha_t + \epsilon \tag{4}$$

Equation 2 outlines the regression for trade from China to the United States, Equation 3 analyses trade from the intermediates to the United States, and Equation 4 examines trade from China to the intermediaries. The TSavings variable was omitted from Equation 2 as this measure makes no sense in the context of this trade flow.

A comparison of the results of each of these regressions will show both transshipment and the extent to which China and the United States are increasing their trade with intermediaries as a result of the trade war. If, for instance, tariff savings account for China increasing exports of j to country i by X amount while also accounting for country i increasing exports of j to the United States by $x + 1$, it can be inferred that American importers are sourcing 1 unit of j from country i while x is being redirected (assuming we see a corresponding decrease in bilateral U.S.-China trade). Conversely, if Chinese exports to i increase by more than country i's exports to the United States, we can deduce that the difference represents a shift in Chinese exports to less expensive destinations. These shifts in trade patterns can then be compared to the changes in direct trade between the United States and China so as to determine the significance of transshipment. Lastly, it should be noted that the comparison strategy necessitates that the regression results be in absolute dollar terms.

4 Data

The data used for this study were collected using the Unites States Census Bureau and the World Integrated Trade Solutions (WITS) tool which aggregates data from UNCTAD, the International Trade Center, the United Nations Statistical Division, and the WTO. Specific data on individual countries' institutional qualities and trade capacities were sourced directly from the World Bank Development Indicators databank.

The dataset spans the years from 2016 through 2018 and contains information on the United States, China, Brunei, Hong Kong, Indonesia, Japan, Korea, the Philippines, Singapore, and Vietnam. The geographical scope is limited to East Asia as the prevailing literature on transshipment strongly suggests that such activities are most likely to be carried out using intermediaries that are close to the exporter in terms of geography, culture, international relations, and trade relationships (Kokko et al., 2007; Jacks and Novy, 2019). Cambodia, Malaysia, Myanmar, Taiwan, and Thailand were omitted for lack of data. Laos was left out due to its landlocked status.[1] While I was able to collect data for the United States through September of 2019, year-to-date information of the other countries is not accessible to me at this time. This shortcoming may skew my results. Any changes in transshipment that may occur as the trade war progresses are likely to be left out as I am only able to analyze trade shifts that occurred within the first year of the dispute. However, the fact that the trade war started in March of 2018 should make the available data sufficiently informative.

Concerning tariff and product selection, this paper utilizes the HS 6-digit classification system across all goods[2] and computes tariffs using a weighted average. The product nomenclature and aggregation level were selected because they represent the most granular measurement that is common to all countries. Because individual reporters each use their own systems for identifying and taxing goods below the 6-digit level, some tariff aggregation had to be done. WITS data could not be used for the United States because its tariffs on China fall outside of the rules that are imposed by international bodies. Since WITS pulls data from these same organizations, it only provides information on the types tariffs that these institutions consider to be acceptable. This is why the American data were sourced from the Census.

The tariffs that the United States imposes on China, and indeed on every other country considered in this paper, were obtained by dividing the calculated duty (the dollar value of

[1]Data on both trade and tariffs was unavailable for these countries. Their exclusion from the sample may understate the results if they are being used for transshipment. If no goods are being rerouted through them, their exclusion should not have any impact on this study.

[2]These were later aggregated to the 2-digit level to make the regressions more manageable.

tariffs collected in a given year for a given import) by the dutiable value (which the census defines as the customs value of imported goods subject to duties). This effectively yields a weighted average of all the tariffs imposed on the subgroups of each HS 6-digit product and is the most accurate possible measure of the real rate applied to each good.

Since the tariffs imposed by the intermediary countries on China theoretically comply with WTO rules, the use of WITS data is appropriate. The specific duty type selected was the Effectively Applied Rate. WITS determines this by analyzing the various rates available for a given good under a variety of tariff schedules; the lowest of these is used to establish the Effectively Applied Rate. One potential limitation of this method is the assumption that exporters will always elect to use the lowest tariff. This may not be an accurate reflection of reality when there are significant administrative or quality control costs associated with qualifying for the lowest possible rate. This being said, the Effectively Applied Rate is the closest approximation of what exporters actually pay. It is also the tariff type that has the least amount of unreported values in WITS. Ad-valorem equivalents (using the UNCTAD method) were additionally factored into the weighted averages for each country. When possible, the WITS data were used to fill in missing values in the American Census data. These account for 2,278 observations or 1.3% of the total for the United States.

There is also missing data on Vietnamese and Philippine imports from China and on the tariffs applied to these in 2018. Imports were therefore substituted by Chinese exports. Theoretically, this should not impact my results as imports and exports must, by definition, be equal to each other. The missing 2018 tariffs were linearly interpolated from the 2016 and 2017 observations. This should not adversely impact my results since Vietnam and the Philippines are not currently engaged in a trade war with China. Therefore, one should not expect the tariffs to suddenly jump in 2018.

Countries' trade capacities were measured using the Quality of Port Infrastructure series from the World Bank Development Indicators databank. This metric is based off of business leaders' views of their own country's port infrastructure on a scale from 1 (extremely undeveloped) to 7 (well developed and efficient by international standards). The average port quality for the sample was 4.91 in 2019 with Singapore having the highest quality (6.67) and the Philippines having the lowest (2.93). The underlying data are collected from the World Economic Forum's Executive Opinions Survey. Missing data points for Brunei were calculated using a ten-year trend. While I believe this measure to be more than adequate, it must be noted that a better alternative exists. The Logistics Performance Index operates in much the same way as the Quality of Port Infrastructure except that it examines all trade and transport related infrastructure and is measured using the views of logistics experts that need not reside in the country under consideration. This is better because it has a broader scope, is less susceptible to domestic biases, and includes foreigner's perceptions which are more useful for the context in which the capacity variable is being used. Unfortunately, this measure contained so many gaps that even the ten-year trend method was unable to produce trustworthy information.

The Institutions variable is measured using the Ease of Doing Business Index. This statistic is gathered from lawyers, freight forwarders, business people, and other professionals who regularly deal with their respective country's legal or regulatory system. These people respond to a standardized survey, that is administered by the World Bank, which is used to measure the degree to which each economy's institutions are conducive to business. This was the only measure of institutions that contained data on all the countries in the sample.

Table 1 shows the average tariff savings across all goods for intermediary countries from 2016 through 2018. From this we can already observe a general trend of tariff savings increasing after the onset of the trade war in 2018. Going back to the Model section of this paper, we can remember that this variable represents the difference in total tariffs paid or avoided when transshipping a good through a third country. Negative values mean that it is most profitable to ship directly from China to the United States. The information presented in Table 1 clearly reflects the shift in American policy from free trade to protectionism with respect to China. Japan and Korea in particular stand out as they have negative tariff savings in all years, even after the onset of the trade war (it is worth noting, however, that these values do become less negative over time). This makes sense when considering the close political and institutional ties between these countries and the United States which should, in theory, mean that they are less likely to be used as illegal transshipment channels. This hypothesis is in line with Kokko et al.'s (2014) findings that these types of relationships impact trade. On the flip side of things, Hong Kong and Singapore had the highest average levels of tariff savings in 2018. This is logical since Singapore is an extremely trade dependent economy with which the United States has preferential agreements and Hong Kong has long been used by China for transshipment (Fishman et al., 2008).

The driving forces behind the numbers in table 1 can be seen in Figure 1. This graph plots the average tariffs imposed by the U.S. on each partner (solid lines) as well as the tariffs that each of these partners impose on China (dashed lines) over time. From this, we can see that the U.S.'s tariffs on China have historically been lower than those which it imposed on the other countries in the sample. Meanwhile, the Asian intermediaries' tariffs on China tend to be the lowest that I observed. Another trend in Figure 1 is the fact that American tariffs (solid lines) have been rising much more than those of the East Asian countries. Together, these trends wound indicate that the tariff savings observed in Table 1 stem more from the U.S. raising trade barriers than from the intermediaries lowering them. Lastly, we can see that the negative savings that are associated Japan and Korea in Table 1 neatly translate into these countries' higher average tariffs on China that we see in Figure 1.

Table 2 depicts the total trade between each intermediate country and the United States. It is immediately apparent that China's exports have continued to grow despite the trade war, albeit at a slower rate. The two countries with the lowest tariff savings, Japan and Korea, saw relatively little change in their outflows before and after the trade war. Interestingly enough, Hong Kong's exports actually decreased in 2018, suggesting that it is not actually being used for transshipment.

5 Results

I begin this section by presenting the results for my regressions as they are outlined in Section III. The shortcomings associated with the results of these lead me to run two more sets of regressions. In the first, I rework the equations to include country fixed effects which are more flexible measures of country-specific differences. I keep these in the second set of new regressions while also redefining the tariff savings variable. These last results represent those of my preferred model and are used to calculate my final estimation of transshipment in the U.S.- China trade war.

The initial regression results are outlined in Table 3. It is encouraging to see that tariffs have a negative impact on trade that is statistically significant at the one percent and ten percent levels for the trade flows between the United States and China and between China and the intermediaries respectively. The beta on tariffs for American imports from the intermediaries had the opposite sign of what was expected. This coefficient, however, is also not statistically significant at any of the conventional levels and is thus not overly concerning.

Regarding the coefficients on tariff savings, these are correctly signed and insignificant for trade between the United States and the intermediaries and are incorrectly signed but significant at the ten percent level for flows from China to the intermediaries. Savings were omitted from the regression in column one of Table 3 because it makes no sense to use this measure for flows that do not pass through any intermediaries. These unexpected results may be caused by the possibility of their being some correlation between tariff savings and certain country-specific factors that are not present in the model. These will be addressed shortly.

The control variables in this regression were, for the most part, correctly signed but insignificant. The only exceptions to this were institutions for the U.S.-China trade flow and capacity for the intermediary-China flow which were both significant at the one percent level. Oddly enough, the sign on the first of these coefficients is negative suggesting that better institutions decrease exports from China to the United States. While this finding is counterintuitive, it is worth remembering that institutional quality is being proxied with the Ease of Doing Business Index. It is possible that this is an imperfect measure that is actually capturing information outside of what it is intended to such as domestic regulations that do not pertain to international trade. It is also possible that the institutions variable is actually meaningless for all trade flows. In this case, the significance that is observed in the U.S.-China trade would be attributed to the fact that this flow only covers two countries while the other regressions include the pool of eight intermediaries. The smoothing out of variation that stems from these is the most likely culprit for the lack of significance of these coefficients.

For the sake of robustness, a second set of regressions was run in which the controls were replaced with country fixed effects. I expect these to yield better results since country fixed effects are more flexible than individual control variables and should thus capture more information. This flexibility comes from the fixed effects' ability to capture all time invariant changes between countries rather than just the two parameters. The institutions and capacity variables were omitted as these are time invariant in the short-term and would thus be collinear with the country fixed effect. The results of these new regressions are outlined in Table 4. A comparison of Tables 3 and 4shows that the coefficients of interest are largely the same in terms of overall significance and signage. More interestingly, the adjusted R squared values for the second and third regressions have each risen by 0.01. While this is admittedly a very small increase, after which the adjusted R squared values were still quite low, it is still evidence in favor of replacing the control variables with the more accurate country fixed effects.

Unfortunately, the tariff savings coefficient for the China-intermediary trade flow are incorrectly signed and statistically insignificant. I believe that this is due to issues with the underlying data. As was previously discussed in Section IV, the tariffs used for this study were the effectively applied rates. These are calculated by selecting the lowest available rate for each product. In theory, exporters would intentionally make their products comply with the criteria necessary to qualify for certain preferential schedules. However, it is possible that, in reality, many exporters find that the cost of qualifying for these special rates exceeds the cost incurred by exporting under a more lenient yet expensive schedule. The American tariffs were calculated using the real amount paid to customs and are thus immune to the possibility of their not being accurate reflections of reality.

Therefore, another robustness check was performed by rerunning the last set of regressions with an altered version of tariff savings. This new explanatory variable, which carries the name U.S. Savings, was calculated by subtracting the American tariffs on the intermediaries from the American tariffs on China for each product during each period. This new measure is an improvement over the original as it avoids any of the previously mentioned issues that may be present in the data on duties imposed by the intermediaries on China. Country fixed effects were included as they were demonstrated to yield superior results. The outcomes of these regressions are detailed in Table 5 and represent my preferred model specification which will be used to obtain my final results.

In terms of statistical significance, signage, and overall explanatory power, the results shown in Tables 4 and 5 are similar. The most significant difference lies in the size of the betas on savings for American imports from the intermediaries. The original coefficients (from Table 4) indicate that trade between the United States and the intermediaries will increase by \$187,287 if overall tariff savings increase by one percentage point. The results of the new regression (Table 5) show that a one percentage point increase in U.S. Savings will raise the same trade flow by \$1,115,326 for the average country and HS 2-digit commodity. The implication is that exports from the pool of eight intermediaries to the United States are motivated more by the tariffs imposed by the United States on these countries than they are by the tariffs that these intermediaries impose on China.

The coefficient for U.S. Savings for trade between the intermediaries and China is just as incorrectly signed and insignificant as it was previously. I believe that this due to China intentionally misreporting its exports to the intermediaries. Transshipment-induced fluctuations in trade between these countries and china is being obfuscated by misreporting on the part of the Chinese. This hypothesis is supported by research conducted by Fishman and Wei (2004) in which they measured Chinese tariff evasion using discrepancies between data reported by Hong Kong versus that which was reported by China. The authors found that this evasion was, in part, caused by China underreporting the value of its exports and misreporting the types of goods it exported in order to qualify for a lower rate. Should such activities be taking place, they would disproportionately influence the results associated with Vietnam and the Philippines as import data for these countries were replaced with exports from China.

The final steps in this study of transshipment were carried out using the coefficients that were obtained in the last set or regressions as these are the most robust and informative. Unfortunately, this portion of the analysis had to be restricted to only the exports from China and the intermediaries to the United States due to the lack of significance and incorrect signage on the savings coefficient in the third column of Table 5. However, the final findings for these flows indicate that imports from China to the United States decreased by $924,716,583 (0.17% of 2018 trade) because of the tariffs while imports from the intermediaries increased by $795,232,230 (0.23% of 2018 trade with the intermediaries, 0.15% of 2018 trade with China) as a result of the tariff savings that were calculated using the second method. These findings are significant to the one and ten percent levels respectively. The equations that were used to obtain the figures above are the following:

$$AverageUSSavings \cdot 1,115.326 \cdot \#Intermediaries \cdot \#HS2Products \cdot 1,000 \tag{5}$$

$$\Delta Tariffs2018 - 2017 \cdot -2,433.207 \cdot \#HS2Products \cdot 1,000 \tag{6}$$

Equation 5 calculates the final result for the U.S.-intermediary flow and Equation 6 estimates those for U.S.-China trade.

The fact that trade between the United States and the intermediaries was motivated by the savings – which are partially defined by the American duties on China – shows that trade between two countries can be, and is, motivated by the tariffs that one of these countries imposes on a completely different partner. This relates to the idea of "multilateral resistance" that was developed by Anderson and van Wincoop (2003). They posit that trade between two countries, country 1 and country 2, will increase when trade barriers between them and the rest of the world are higher than the barriers they have between each other. This would make the price of goods from country 1 in country 2 cheaper than those imported from elsewhere. This, in turn, would increase demand for these goods in country 2, resulting in higher trade. In the context of this paper, one might consider than the increased trade barriers between the United States and China would increase trade between the United States and other countries. Therefore, transshipment can be seen as China piggybacking off of these new relationships to circumvent the negative impacts of the trade war.

The magnitude of my results is far below what was expected based off of the literature. This is especially true when comparing the fact that transshipment only accounted for 0.15% of U.S. imports from China in 2018 with Fung's (1996) discovery that the U.S.-China trade deficit was being understated by 35% due to such activities in the 1990s. It is worth keeping in mind that China's relation with the rest of the world has changed since Fung's paper which was written before China joined the WTO. Furthermore, these results are not exact measures but are rather aggregates of the impacts on trade of the average commodity from the average intermediary, as can be seen from equations 5 and 6. This, combined with the fact that several potential intermediaries were lost due to lack of data and the possibility that China is misreporting its trade, leads me to believe that my results are understated compared to reality. Nevertheless, it is clear that close to 86% of the imports from China that are being lost due to the tariffs are being compensated for by increased imports from the intermediaries due to the tariff savings associated with them.

Because none of the analyses in this paper were able to find reliable conclusions regarding the trade between China and the intermediaries, it is impossible to say whether or not the increase in imports from the intermediaries represents transshipment or direct sourcing from these countries. However, an examination of Tables 3, 4, and 5 shows that the coefficients on straight tariffs for this trade flow are insignificant in all cases while savings are always significant at the ten percent level when including country fixed effects. If the increase in trade was solely due to direct sourcing, one would expect the tariffs that the United States imposes on these countries to be the principle motivator. Since I found that this shift was instead caused by the tariff savings associated with rerouting trade through the intermediaries, it is more likely that goods are being transshipped through the pool of eight countries under consideration.

High levels of transshipment should imply that trade wars are ineffective foreign policy tools as such activities mitigate the impact of these actions. However, the finding that only 0.15% of U.S.-China imports were rerouted means that most of the impacts of the trade war are still being felt. While it is likely that these results are understated, it will take a significantly higher amount of transshipment to draw any sound policy implications.

6 Conclusion

This paper seeks to both identify and quantify transshipment in U.S.-China trade war.

This research was motivated by evidence of China's past transshipping activities and numerous media reports of such activities happening again in the context of the ongoing trade war. By comparing tariff induced changes in the trade flows between the United States and China, China and a pool of intermediaries, and the intermediaries and the United States, I find statistically significant evidence (at the ten percent level) that transshipment occurred in 2018. More specifically, $795,232,230 were rerouted from China, through other East Asian countries, to the United States. This figure accounts for approximately 0.15% of American imports from China in 2018. These results are most likely understated due to this paper's data limitations and aggregation techniques. Nevertheless, they do serve as a baseline which could be used to inform other works on transshipment in the U.S-China trade war.

This paper's findings leave much room for future research. Repeating this research with consistent, up to date data that includes a more comprehensive pool of intermediaries would go a long way towards establishing a more precise measurement of transshipment. Furthermore, estimation using an alternative technique, perhaps one that compares trade statistics when reported by different countries, might be useful in eliminating possible obfuscation via misreporting. Lastly, it could be interesting to break down transshipment by country and product so as to gain a more detailed understanding of this behavior.

References

Abaid, A. e. a. (2018). The impact of trade conflicts on developing Asia. December Asian Development Bank Working Papers Series.

Amiti, M., Redding, S. J., and Weinstein, D. (2019). The impact of the 2018 trade war on U.S. prices and welfare. NBER Working Paper no. 25672.

Anderson, J. E. and van Wincoop, E. (2003). Gravity and gravitas: A solution to the border puzzle. *American Economic Review*, 93(1):170–192.

Fishman, R., Moustakerski, P., and Wei, S.-J. (2008). Outsourcing tariff evasion: A new explanation for entrepot trade. *The Review of Economics and Statistics*, 90(3):587–592.

Fishman, R. and Wei, S.-J. (2004). Tax rates and tax evasion: Evidence from "missing imports" in China. *Journal of Political Economy*, 112(2):471–496.

Fung, K. C. (1998). Accounting for Chinese trade: Some national and regional considerations. In Baldwin, R. E., Lipsey, R. E., and Richardson, J. D., editors, *Geography and Ownership as Bases for Economic Accounting*. University of Chicago Press, Chicago.

Ishimine, T. (1979). Organization, ideology, and performance of China's international trade. *Pakistan Economic and Social Review*, 17(1-2):1–22.

Jacks, D. S. and Novy, D. (2019). Trade blocks and trade wars during the interwar period.

Kokko, A., Soderlund, B., and Tingvall, P. G. (2014). Redirecting international trade: Contracts, conflicts, and institutions. *Journal of Economics and Statistics*, 234(6):688–721.

Tan, D. and Chen, C. (2019). Modeling the economic impact of the Sino-US trade dispute: A global perspective. In Song, L., Zhou, Y., and Hurst, L., editors, *The Chinese Economic Transformation: Views from Young Economists*. ANU Press, Canberra.

Figures

Figure 1: Average Tariffs Between the U.S. and Partners and Partners and China

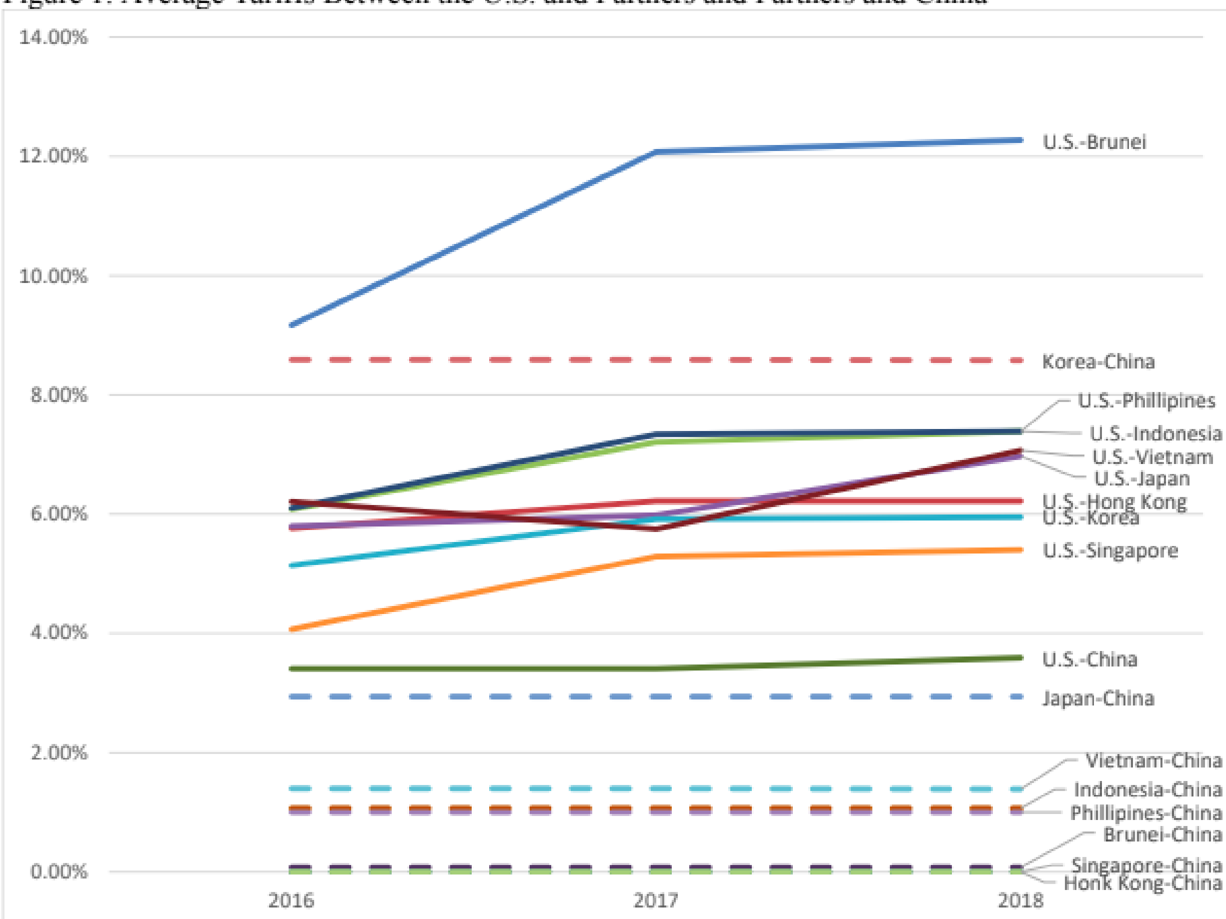


Table 1: Average Tariff Savings (all goods, in percentage points)

Country	**2016**	**2017**	**2018**
Brunei	-0.04%	-0.5%	0.5%
Hong Kong	0.05%	0.09%	2.52%
Indonesia	-1.52%	-1.98%	1.39%
Japan	-3.37%	-3.40%	-0.58%
Korea	-11.26%	-7.71%	-4.80%
Singapore	1.17%	-0.08%	3.58%
Philippines	-2.16%	-2.63%	0.02%
Vietnam	-1.55%	-1.43%	1.05%

Tables

Table 2: Total Imports/Exports to the United States (all goods)

Country	2016	2017	2018
China	462,461,088 (n/a)	505,731,861 (9%)	540,200,185 (7%)
Brunei	20,730 (n/a)	22,798 (10%)	98,015 (330%)
Hong Kong	7,454,600 (n/a)	7,373,891 (-1%)	6,285,965 (-15%)
Indonesia	19,230,682 (n/a)	20,293,332 (6%)	20,959,655 (3%)
Japan	132,105,835 (n/a)	136,949,218 (4%)	143,193,173 (5%)
Korea	69,978,088 (n/a)	74,307,133 (6%)	77,469,914 (4%)
Philippines	10,054,614 (n/a)	11,637,074 (16%)	12,610,043 (8%)
Singapore	17,865,308 (n/a)	20,159,469 (13%)	27,442,135 (36%)
Vietnam	42,110,250 (n/a)	46,477,086 (10%)	49,175,663 (6%)

Note: values shown in thousands of dollars, year-over-year changes shown in parenthesis

The Effects of the Minimum Wage on Social-Safety-Net Dependence Over Time

Camille Baker

May 5, 2020

Abstract

Although there is much research on the effects of minimum wage increases on workers and low-income families, there is little that investigates how these effects persist or dissipate over time. Using an event-study specification, I investigate how minimum-wage changes affect family incomes at various multiples of poverty, as well as eligibility for and participation in the Supplemental Nutrition Assistance Program (SNAP). I compare the latter effect to that obtained from a state panel regression approach used in previous literature. I find evidence that minimum-wage increases reduce the prevalence of low family income and SNAP participation, but that these effects dissipate by 5 quarters post-increase. At its peak, the effect on SNAP participation is similar in size to that obtained from a state-panel specification.

Acknowledgements

I am grateful to my advisor, Sarah West, and my readers, Samantha Cakir and Leslie Myint; to Amy Damon and my fellow honors seminar participants for additional feedback and guidance; and to many other professors and fellow students for helpful discussions.

1 Introduction and Literature Review

The question of raising the minimum wage is hotly debated in politics today. On its face, the effect of a higher minimum wage is simply to increase the hourly wage of the low-wage workers for whom it is binding. However, some fear that minimum-wage increases may have harmful effects on labor demand, ultimately reducing income and increasing poverty among those the minimum wage is intended to help.

The question of how the minimum wage affects the income distribution is relevant not only in itself, but because of how it may affect the social safety net. Eligibility for social safety net programs is typically based on income, among other factors. Thus, to the extent that increases in the minimum wage boost income for needy households, we should expect them to reduce participation in social safety net programs. If the minimum wage does increase income for low-income households, therefore, the reduction in government expenditures on the social safety net could represent a further benefit of increasing the minimum wage.

A further important consideration regarding this question is the time span of the effects of minimum-wage increases. If minimum-wage increases produce long-term gains in income, and this leads to independence from the safety net, that may be considered cause to prioritize the minimum wage from a policy standpoint. If, however, its effects quickly dissipate – or worse, have long-term harmful effects on income and safety-net dependence by accelerating the automation of low-wage workers' jobs – that could be a serious flaw in the minimum wage as a policy.

Despite the importance of the potential dissipation of effects over time, there is little research on the impact of minimum-wage increases over time – and none that focuses on the income distribution or social safety net programs. To fill this gap, I use an event-study approach to examine the effects of a minimum wage increase over time in the period following the increase.

In a typical event study at the state level, the level of the dependent variable in a state is measured at each time increment before and after that state experienced an event (in this case, a minimum wage increase). These trends are combined for states that experienced events at different times (with event times aligned such that, for example, states 1 quarter post-increase are all considered together), and compared to those that did not experience an event. This helps to eliminate omitted variable bias that may occur if we study time trends within a single state. In the case of the minimum wage, one obstacle to the use of event studies is that multiple minimum-wage increases may occur in the same state or region in quick succession. To overcome this, I use Sandler and Sandler's (2014) multiple event study methodology (described in more detail below) to account for the possibility of multiple minimum-wage changes in a single state. This allows me to take advantage of all changes in the binding minimum wage of one or more states, whether from a state- or federal-level law change.

I further examine how changes in the income distribution may affect social safety net program participation by running the same specifications to measure the effect on participation in one particular safety-net program. Because of its widespread usage and broad eligibility standards, my chosen program is the Supplemental Nutrition Assistance Program (SNAP). I find that, while increasing the minimum wage reduces poverty and near-poverty for about a year after the change, it does not affect SNAP participation.

1.1 Overview of SNAP

SNAP, historically known as food stamps, is a federally-funded, state-administered program that provides households with benefits that they can spend on most kinds of food at grocery stores. In 2017, roughly 9.2% of US households were on SNAP for at least part of the year; these households included about 16.7% of US children. On average, each household received about $250 in monthly benefits from the program (Watson, 2019).

A household is eligible for SNAP if its income is below a certain cutoff specific to the number of people in the household. Households containing an elderly or disabled person face a more inclusive cutoff. These income limits are set each year by the federal government at 130% of the poverty line. However, states can set higher limits (up to 200% of the poverty line) through a process called broad-based categorical eligibility (BBCE), where households are automatically eligible for SNAP if they qualify for other designated programs with higher income limits. BBCE can also be applied to other state-level programs aimed at specific populations. For example, in Minnesota, the income limit for SNAP is set to 165% of the poverty line (Table 1a); in the 40 total states that use BBCE, income limits range from the minimum 130% to the maximum 200% (Table 1b).

Households eligible for SNAP must submit an application to the state government, and interview by phone or in person with a case worker who instructs applicants on supplying the proper documents to prove they qualify. Once accepted, the amount of benefits a household receives is based on its income; the closer it is to the cutoff, the lower the benefit amount. Households must re-apply for benefits each year.

1.2 Overview of Minimum Wage Effects

The most basic and direct impact of the minimum wage is to increase wages for those earning below it – thereby increasing income for them and their families. However, the minimum wage may also affect the markets for low-wage labor and the goods and services it produces. This may produce downsides: firms reducing the number of workers or labor hours they hire; lower corporate profits; or increases in the prices of goods produced by minimum-wage workers. In particular, a reduction in labor hours or employment may counteract some or even all of of the income-boosting effects of the wage increase itself; or cause some families to lose income while others gain it. Thus it is not clear *a priori* how many families will move out of poverty or SNAP eligibility in the aftermath of a minimum-wage increase; or how many others may move into these categories instead.

Another incentive that may change in response to a minimum-wage increase is on the labor supply side. A concern in designing social-safety-net programs is the possibility of creating a "welfare cliff". That is, if gaining a little bit of earnings decreases benefits received from social-safety-net programs by a large amount, it may actually decrease overall income, incentivizing recipients to keep their earnings low (potentially by working fewer hours). The federal eligibility requirements for SNAP include the requirement that SNAP recipients not voluntarily reduce their working hours. However, SNAP recipients affected by the welfare cliff may keep their hours lower through other means. For example, a worker may not seek out additional hours or shifts when the opportunity arises; may not look for a new job if their current employer cuts their hours; or may seek fewer hours when starting a new job. On the other hand, since SNAP benefits decrease steadily as household income approaches the maximum cutoff, rather than abruptly dropping to zero, this effect may not play a meaningful role in the case of SNAP.

Because of these possible factors, the effect of a minimum-wage increase on poverty or SNAP may not reflect the simple "mechanical" effect of the increase – that is, the change in poverty or SNAP that we would predict if we increased the hourly wages of minimum-wage workers while holding all else constant, including hours worked. In real life, hours worked may change in response to a minimum-wage increase, affecting total income. Moreover, not everyone who is eligible for SNAP receives it – some may be unaware that they have moved into eligibility, or the application process may act as a burden. Thus even a change in the income distribution in response to the minimum wage may not fully translate into a corresponding change in SNAP participation. Understanding these effects can help us evaluate minimum wage increases both in terms of the welfare impact on people at the margin of SNAP eligibility, and in terms of money potentially saved by the government on SNAP.

1.3 Employment, Hours, and Wage Spillover Effects of the Minimum Wage

Economic theory predicts that, if the labor market is competitive, imposing a binding minimum wage will decrease firms' demand for labor. If, however, firms have monopsony power in the labor market, it is possible to construct a theoretical framework in which a binding minimum wage will not reduce labor demand. Much of the empirical research on minimum wage laws, therefore, focuses on which of these scenarios more closely matches real-world markets – i.e., whether the minimum wage causes employers to reduce the labor they hire (e.g., Neumark et al. (2004); Dube et al. (2010)). Other research focuses on related questions, such as the minimum wage's impact on poverty and the income distribution (e.g., Addison and Blackburn (1999); Dube (2019)).

Early empirical minimum-wage research relied mostly on time-series methods, regressing a dependent variable (e.g., employment) on the national-level minimum wage before and after a change while controlling for macroeconomic trends. Neumark and Wascher (1992) provide a full review of this early literature. Time-series methods like this are used occasionally in more recent studies (e.g., Wolfson and Belman (2004)), but since the 1990s, researchers have mostly focused on more sophisticated strategies exploiting variation across space as well as time. One early example is Card and Krueger (1994), who use a "case study" difference-in-differences approach, comparing employment trends in nearby regions in New Jersey and Pennsylvania before and after New Jersey's minimum wage increased (while Pennsylvania's remained constant). They find no significant effect. Similar results are found by Card (1992), who uses a similar case-study approach to study the minimum wage's effect on teen employment by comparing California (whose minimum wage increased) to a group of other states. Neumark and Wascher (1992) provides an early example of another now-common approach, the state-level panel regression with time and state fixed effects. They find that the minimum wage has a negative impact on hours worked (not people employed).

Since then, many studies have examined the employment and hours effects of the minimum wage using many strategies, and have found mixed results. Most studies find negative employment and hours elasticities with respect to the minimum wage. However, these are typically between 0 and -0.5, suggesting that reductions in employment and hours do not fully cancel out wage increases. Some studies (Neumark et al. (2004); Clemens and Strain

(2017)) exploit multiple state-level minimum wage changes to run difference-in-differences specifications with minimum-wage increases as treatments. Neumark et al. (2004) find employment elasticities of -0.12 to -0.17 immediately following a minimum-wage increase; they find no initial effect on hours conditional on remaining employed, but an elasticity of -0.2 to -0.25 one year later. Dube et al. (2010), meanwhile, obtain an overall labor demand elasticity of about -0.48. Sabia (2009) estimates a state-level panel regression with fixed effects; he finds that a 10% increase in the minimum wage leads to a 1% average decrease in both employment and hours. Clemens and Strain (2017) estimate employment effects for various age-and-education groups, finding employment reductions ranging from zero to about 2 percentage points in response to typical minimum-wage increases. Cengiz et al. (2019), however, using a bunching estimator approach based on the number of jobs above and below a new minimum wage, find no significant employment effects.

Some research disaggregates the employment and hours effects of the minimum wage by worker group. These studies generally find that those workers with the least education and experience see the greatest reductions in employment and hours. This is of particular importance in evaluating the welfare effects of the minimum wage, since these workers may be more or less likely to account for a large percentage of their families' income. Neumark et al. (2004) find that the employment and hours elasticities of the minimum wage vary widely across groups, and for the least educated groups, may have magnitude greater than 1 (indicating a reduction in earned income). Clemens and Strain (2017) run a triple-differences specification of the minimum wage's effects on employment and hours over skill level, and find that lower education and experience are associated with larger effects. In particular, the typical minimum-wage increase reduces employment by 1.6 to 2.1 percentage points for those under 25 without a high school diploma. Sabia (2008) finds no effect of a minimum wage increase on hours for single mothers in general, but for the least educated ones, a significant negative effect on hours that fully cancels out the increase in wages. A later study by Sabia (2009) finds an overall hours elasticity of -0.1 with respect to the minimum wage, but finds that this effect is largely driven by the effect on the least experienced workers. Similarly, in Meer and West's (2016) study of minimum wage increases and slowing job growth, they find that growth slows most dramatically in employment of younger workers and in low-wage industries.

The effects of the minimum wage extend beyond just minimum-wage earners. Many studies (Neumark et al. (2004); Cengiz et al. (2019)) find evidence of a spillover effect: when the minimum wage increases, the wages of those earning slightly more than the previous minimum wage also increase. Neumark et al. (2004) find that the spillover effect extends only to those near the minimum wage; higher earners' wages are not affected. Phelan (2019) also finds evidence of a spillover effect, and proposes a mechanism: a minimum wage increase is equivalent to a decrease in the compensating differentials for more undesirable jobs available to minimum-wage workers, reducing labor supply for those jobs and causing equilibrium wages to rise.

1.4 Effects on Income, Poverty, and Program Participation

The extent to which the effects of the minimum wage on individual earners will affect families depends whether affected workers' earnings represent a large portion of their families' income. If many minimum-wage earners are part-time teen workers whose parents earn substantially more money than they do, the effect on families may be negligible. However, the literature shows that these groups form the minority of all minimum-wage earners. 56% of workers earning at most the minimum wage are adults over 25, along with 60% of those earning at most 1.25 times the minimum wage. 9% of both groups are single mothers (Belman et al., 2015). One in every three minimum-wage workers is the only worker in their household (Leigh, 2007). This suggests that family earnings and income could be meaningfully impacted by minimum-wage increases.

Given the wide range of findings on employment and hours effects, it is not obvious how – and by how much – the minimum wage should affect income and poverty. If the employment elasticity for a particular worker group is -1 (that is, quantity scales down by exactly as much as wage scales up), then an increase in the minimum wage will have zero effect on overall earned income for this worker group, since the reduction in hours will exactly cancel out the higher wage. However, if this effect is primarily in the form of an employment effect (some workers lose their jobs and the remaining ones keep the same number of hours), it could nevertheless affect the poverty rate if those who lose their jobs enter poverty while those whose income increases stay out of it. These effects may be smaller or larger depending on if the employment elasticity with respect to the minimum wage is elastic or inelastic.

Compared to employment and hours effects, there are relatively few empirical studies that directly address the effect of the minimum wage on income, poverty, and the overall income distribution. Addison and Blackburn (1999) note that most previous work simply simulates the effect of the minimum wage on poverty by evaluating the effect of increasing workers' wages, without taking into account the possibility of other labor-market effects that may affect income. They run a state-level panel regression on logged minimum wage with time and state fixed effects, and find that raising the minimum wage by 10% reduces poverty by 5%; this result holds for teens as well as adults with low education.

Many other studies, however, have found less encouraging results. Vedder and Gallaway (2002) use a state panel approach, controlling for macroeconomic variables and federal transfers, to estimate minimum-wage effects on poverty for a number of sub-populations as well as the population as a whole. They find no significant effects in either direction. Leigh (2008) simulates the effect of a minimum wage increase on income, but takes into account employment effects by taking plausible numbers from existing literature; he also finds no significant effect. Neumark and Wascher (2002) provide a more detailed examination of these effects, using CPS microdata to estimate a model of individuals' probability of moving from poor to non-poor status or vice versa. They find that increasing the minimum wage produces about the same amount of movement in each direction, producing a zero effect in aggregate (although incomes do increase for those who remain poor). In a later study, Neumark et al. (2005) use CPS data following families over consecutive pairs of years, and run a difference-in-differences specification comparing those who lived in states that saw a minimum-wage increase to those in other states; they find no general trend toward increasing or decreasing incomes for treated households. They also run a state-level version of this specification studying changes in the fraction of people below a certain multiple of poverty. They find that minimum-wage states have more people overall below poverty in year 2, but fewer below 50% of poverty.

These results are contradicted by a recent paper by Dube (2019), who runs several state panel regressions of the proportion of families at several different multiples of the poverty line, with leads and lags of the minimum wage as well as time and state fixed effects and state-specific linear trends. Additionally, he runs an unconditional quantile regression of the income distribution using that same specification. He finds that the elasticity of the non-elderly poverty rate with respect to the minimum wage is between -0.22 and -0.46, and the elasticity of the tenth and fifteenth quantiles of family income are between 0.15 and 0.43, indicating that raising the minimum wage reduces poverty while boosting the tenth and fifteenth quantiles of income.

There is very little literature on the effects of the minimum wage on social safety net programs for low-income people. Only one paper (Reich and West, 2015) focuses on the impact of minimum wage on SNAP. Using a standard state panel approach including Census-division-specific year fixed effects, they find that a 10% increase in the minimum wage is associated with a 2.4-3.2% reduction in SNAP enrollment.

1.5 Event Study Methodology

The event study approach, which I use in this study, has been rarely used in existing minimum-wage literature. An NBER working paper by Adams et al. (2018) provides one of the only examples of a minimum-wage study using an event-study specification with a full set of pre- and post-event dummies. In this paper, they study the effect of minimum-wage increases on labor-market search effort, using individual-level panel data. Their specification includes dummy variables representing, for each observation in their data, whether it is n months from a month in which the minimum wage goes up, where n ranges from -5 to 5 (the largest number they could choose without allowing event windows to overlap within states). They also include month and state-month fixed effects. They find large positive effects on search intensity immediately following an increase, but find that these effects quickly dissipate after the first month post-event.

One major advantage of an event-study specification in minimum wage research is that some literature points to the importance of the time scale over which the effects of the minimum wage take place. Neumark et al. (2004) find negative employment effects immediately following increases in the minimum wage, but find effects on hours only at a one-year lag. Moreover, they find that individual earned income increases immediately after an increase, but has decreased by a year later. Meer and West (2016), in a local-level panel regression studying employment effects of the minimum wage, find that the effects lessen over time but are still present as far as eight years post-increase.

Another benefit of event studies is that they provide a natural way of checking for pre-trends; that is, the possibility that one's dependent variable tends to trend upward or downward in the time leading up to events (indicating potential endogeneity). Multiple studies (Dube et al., 2010; Allegretto et al., 2017) find evidence that this may bias results of minimum wage studies. Dube et al. (2010) address this by using a difference-in-differences estimator on border-county pairs; they find a labor demand elasticity of about -0.48, compared to an elasticity of -1 obtained from a traditional state panel regression. Event studies provide another way of addressing this issue (by yielding coefficients for the time periods leading up to events), while also providing information about effects over time.

1.6 Literature Gap

In this paper, I provide the first evidence on the effects of minimum-wage increases on the income distribution as those effects progress over time, rather than in a single period relative to the minimum wage increase. I also provide a link between the literature on income and poverty effects of the minimum wage and its effects on social safety net participation, by evaluating its effects on income at the margin of SNAP eligibility.

Measuring the progress of effects over time is important both from a policy perspective and an empirical measurement perspective. The benefits and costs of the minimum wage depend heavily on whether it produces long-term income increases for those at the low end of the income distribution, or whether it produces only immediate benefits that quickly fade (or become harmful). It is plausible, moreover, that the magnitude of the effect may differ depending on the time frame we choose to study. Labor markets may take time to adjust to the imposition of a minimum wage, as firms make and implement different decisions about factors of production to use. Thus measuring the contemporaneous effects of MW increases, or the effects lagged by a particular amount, may not fully capture a full picture of the effects. In this paper, I provide evidence on the effects of a minimum wage increase as they unfold over time.

Finally, my method allows for comparison between the effects of an increase on the income distribution and the effects on SNAP participation (which should theoretically be linked). Most previous studies on the minimum wage's effect on the income distribution have not addressed the subsequent effects on social safety net programs like SNAP. Meanwhile, studies that focus on SNAP (e.g., Reich and West, 2015) do not address the underlying changes in the income distribution that may produce the effect they find.

2 Economic Theory

2.1 The Minimum Wage and Workers' Income

My theoretical model begins with a model of how minimum-wage increases may affect workers' income, and thus poverty rates. This is loosely adapted from the model proposed by Fields and Kanbur (2007). They begin with a competitive labor market in which labor is a homogeneous input to production initially provided by identical workers at a market wage $w*$. There is a downward-sloping labor demand curve $D(w)$, defined such that $D(w*) = 1$. Suppose a minimum wage $\hat{w}$ is put into place, and we have $D(\hat{w}) = x$. Since D is downward sloping, we have $x < 1$, and a fraction $1 - x$ of the population is unemployed.

For poverty threshold z, a poverty index P_α is defined as

$$P_\alpha = \frac{1}{n} \sum_{i=1}^{q} \left(\frac{z - y_i}{z} \right)^\alpha$$

where there are n individuals in the population, of whom individuals 1 through q have income below the poverty line; and y_i is the income of individual i. When $\alpha = 0$, this simplifies to the so-called "headcount ratio" used to define the poverty rate in the United States and many other countries.

I expand this model to one including multiple time periods, where the labor market takes one period after a minimum-wage increase to respond to it.

- In period 0, the minimum wage is at w_1, and we have $D(w_1) = 1$. Each individual's earnings are w_1.
- In period 1, the minimum wage is increased to w_2, such that $D(w_2) = x < 1$. However, the labor market does not respond in period 1; the same number of people are employed,

and their earnings are w_2.

- In period 2, the labor market responds, and the quantity of labor demanded is x. Firms reduce labor partially by firing workers and partially by reducing the hours of the workers they keep. The parameter p represents what proportion of that is accomplished through hours reductions: the number of employees is multiplied by x^{1-p}, and hours for the remaining employees are multiplied by x^p.[1]

  This means $1 - x^{1-p}$ of workers' income drops to zero. For the rest, their income becomes w_2x^p. If the elasticity of demand for labor is less than x^{-1+p} at this point, this will represent an increase in income; otherwise, a decrease.

  Thus, depending on the poverty line and elasticity, multiple scenarios are possible. In one scenario, only the $1 - x^{1-p}$ fraction of unemployed workers enter poverty, and the rest face hours reductions that cancel out their increased wages and remain as they are. If labor demand is inelastic or hours reductions for remaining employees are small, the workers who remain employed may be lifted out of poverty. However, if labor demand is elastic and the demand response consists primarily of hours reductions, we may see all low-wage workers enter poverty. There could also be a scenario, for high poverty thresholds, where everyone begins and ends in poverty (or, for low poverty thresholds, where everyone begins and ends above poverty).

- In period 3, inflation progresses to the point that the real minimum wage is once again w_1. If the previous minimum-wage increase caused firms to make permanent changes to

[1] For example, if $p = 0.9$, then the number of employees is multiplied by $x^{0.1}$, which is only slightly less than 1; and hours are multiplied by $x^{0.9}$. Thus total labor hours are reduced to $x^{0.1}x^{0.9} = x$, and the reduction is mostly driven by reducing employees' hours.

production (e.g., incentivized the development of technology that makes more capital-intensive production optimal even now that the higher wage is gone), the quantity of labor demanded may be less than 1, and income may be less than the w_1 it was in period 0. However, if this did not happen, income will return to w_1.

Thus, with respect to a headcount-ratio measure, this model predicts that the effect of the minimum wage on incomes below a given poverty threshold depends on the labor demand response to the higher wage. If most low-wage workers remain employed and those workers see little reduction in their hours, the prevalence of low incomes will decrease on the margin of the incomes employed workers are earning. If many workers lose their jobs, the prevalence of very low or zero incomes will increase; if workers remain employed but with fewer hours, the prevalence of low incomes will increase on the margin of employed workers' incomes. Regardless, after some time, the effect should shrink; and eventually poverty rates should return to their prior levels if the production process has not been changed in the interim.

These predictions should apply to income-based SNAP eligibility rates as well as standard poverty-headcount measures. However, they may not fully extend to SNAP participation itself. The following subsection explains how this may occur.

2.2 Factors in SNAP Participation

A household's decision to participate in SNAP in a period t is a function of three things:

(1) The household's financial situation, including wage earnings (as well as any other income sources, a disabled or elderly household member, or anything else that's taken into account).

(2) The eligibility rules in that state, including where it sets the cutoff (between the required 130% and 200% of poverty), and any other programs it chooses to designate for broad-based categorical eligibility (i.e., anyone eligible for that program would automatically be eligible for SNAP).

(3) The costs of applying for and being approved to receive SNAP (including the opportunity cost of the time and mental energy used to fill out forms, set up a meeting, and collect documentation; as well as the disutility of the application process itself).

Together, (1) and (2) determine if a household has the option of applying for SNAP. Assuming the household is eligible, it decides whether to apply or not based on (3), by comparing the amount of benefits to the perceived time/effort cost.

For a simple example, consider a minimum-wage worker (earning a wage of w_1) in a SNAP-eligible household who works h_1 hours per month. Income from sources other than this worker's earnings (either the earnings of other family members, or nonwage income) is constant at I, meaning total income is $w_1 h_1 + I$.

A household of this size is eligible for SNAP if and only if its monthly income is less than the cutoff X. The household's exempt income (that which is deducted from the total income when determining benefit amount) is constant at E, and the maximum benefit is B_{max}. Thus the benefit amount is $B_{max} - (w_1 h_1 + I - E)$, or $B_{max} + E - (w_1 h_1 + I)$.

If the costs of applying for and receiving SNAP (given that you're eligible) is C, then the household will receive SNAP if and only if

$$w_1 h_1 + I < X \quad \& \quad B_{max} + E - (w_1 h_1 + I) > C$$

$$w_1 h_1 + I < \min[X, B_{max} + E - C]$$

If the minimum wage increases from w_1 to w_2, but hours remain at h_1 and all other variables remain constant, the left-hand side of this inequality will grow. This means that for some households, the inequality will no longer hold, and they will either no longer be eligible for SNAP or no longer find it worth it to apply. Thus SNAP participation will decrease. Moreover, even for households that remain, the benefit amount $B_{max} + E - (w_1 h + I)$ will decrease. Thus, if we assume all other variables remain constant, we predict that a minimum-wage increase will decrease SNAP participation at the extensive level, and reduce benefits for those who remain on the program.

Now consider the possibility that hours will change in response to the minimum-wage increase. Assume that, in the presence of a binding minimum wage, labor demand is the binding constraint on hours worked, even if the worker may wish to work more hours at the new wage. Thus, for small wage changes, the new hours worked are approximately

$$h_2 = h_1 + \epsilon_D^w (w_2 - w_1)\frac{h_1}{w_1}$$

where ϵ_D^w is the own-wage elasticity of labor demand. Therefore, after the minimum-wage increase, the households receiving SNAP will be the ones for which the following inequality holds:

$$w_2\left(h_1 + \epsilon_D^w (w_2 - w_1)\frac{h_1}{w_1}\right) + I < \min[X, B_{max} + E - C]$$

When $\epsilon_D^w = 0$ (perfectly inelastic labor demand), the left-hand side of the inequality equals $w_2 h_1$, and the observed change is the same as the mechanical change. When $\epsilon_D^w = -1$ (unit-elastic labor demand), the left-hand side of the inequality is approximately $w_1 h_1 + I$. That is, things are the same as they were before the minimum wage change, and the behavioral change exactly cancels out the mechanical change, and the observed change is zero. When

$-1 < \epsilon_D^w < 0$ (inelastic labor demand), the behavioral change will be nonzero but smaller than the mechanical change, and the overall observed change will be negative.

The results may also be affected if we consider the possibility of imperfect information. We can account for this by assuming instead that a household will apply for SNAP if and only if

$$w_1 h_1 + I < \min[X', B'_{max} + E' - C']$$

where X', B'_{max}, E', and C' represent the household's beliefs about, respectively, the SNAP cutoff, the maximum benefit, the amount of their income that is exempt when calculating benefit amount, and the cost of applying.

If $B'_{max} > B_{max}$ or $E' > E$, or if $C' < C$, the household may apply for and receive SNAP even if its benefit amount does not justify the cost in terms of utility. If $X' > X$, the household may apply for SNAP thinking it is eligible and be rejected.

Meanwhile, if $B'_{max} < B_{max}$, $E' < E$, or $X' < X$, or if $C' > C$, the household may choose not to apply for SNAP even if doing so would increase its utility. This could be a potential cause of low takeup.

If any of these inequalities are more or less likely to hold for the marginal households who leave SNAP eligibility when the minimum wage increases (relative to eligible people in general), it may result in a change in SNAP participation that is higher or lower than we might expect.

3 Empirical Strategy

The theoretical models above suggest that increasing the minimum wage will increase family incomes at a particular level (in particular, at the margin of SNAP eligibility) if those incomes are supported by minimum-wage workers who face inelastic demand for their labor. However, incomes will stay the same or decrease if labor demand is more responsive to the higher wage. Even if families move out of (or into) SNAP eligibility, moreover, participation may not change if the families affected tend not to take up SNAP. All these effects will tend toward zero over time as inflation cancels out a minimum-wage increase.

To test these predictions, I use an event-study specification to measure the effects of minimum-wage increases, over time, on the fractions of families earning below each of several different income cutoffs. I also use the same specification to evaluate the effects on a simple measure of simulated SNAP eligibility, and on SNAP participation rates. I then compare the results of these regressions to those of a state panel regression replicating that used by Reich and West (2015).

The main specification is as follows:

$$Y_{s,t} = \sum_{q} \beta_q (QuartersFromEvent = q)_{s,t} + \beta_c Controls_{s,t} + \alpha_s + \gamma_q + \epsilon_{s,t}$$

where α_s represents state fixed effects and γ_q represents quarter-within-year fixed effects (i.e., fixed effects for whether a particular quarter is the first, second, third, or fourth of the year). In addition to these fixed effects, I variously control for the year in two different ways: with a linear year term in the control vector, and with year fixed effects; as well as running a version without year effects. (I elaborate more on this below.)

In different versions of this regression, $Y_{s,t}$ stands for various different dependent variables, described in more detail below. The time-to-event dummies $\sum_q \beta_q(QuartersFromEvent = q)_{s,t}$ run from n quarters before the event to n quarters after, for multiple different values of n. These ranges sometimes overlap for different minimum-wage increases in a particular state. To address this, I use the method proposed by Sandler and Sandler (2014) of allowing more than one time-from-event dummy variable to equal 1 for the same state and quarter.

The number of leads and lags around each event is informed by the theoretical prediction that the effects of a minimum-wage increase will dissipate after inflation causes the real minimum wage to return to its pre-increase value. The inflation rate over the period studied was roughly constant at about 2%. The typical minimum wage increase, meanwhile, was about 5-8%; it would take about 3 years for such an increase to be canceled out by inflation. Thus for the longest-time-scale version, I use a maximum of 12 quarters for the time frame.

I run several versions of this regression:

- I use four different ways of defining an event. In one version of the regression (Tables 3a-b), every minimum-wage change of at least 5 percent (relative to the mean minimum wage for that state and quarter) is treated as an event. In two other versions (Tables 4a-4b and 5a-5b), the thresholds are instead 7.5% and 10%. In a final version, following Sandler and Sandler (2014), I include pre- and post-event "dummies" for every minimum-wage change, which are equal to the percent increase in the minimum wage. Because these measures are defined in terms of the existing minimum wage rather than absolute amounts, there is no bias introduced by inflation rendering the definitions less exclusive over the time period studied. (Meanwhile the poverty measures are defined in terms of the federal poverty thresholds, which are adjusted to account for inflation.)

- For each of the four approaches above, I run the regression with seven different dependent variables: the fraction of a state's families at or under 75%, 100%, 125%, 150%, and 200% of the poverty threshold for their year and family size; for the fraction of families estimated to be eligible for SNAP; and for the fraction of families who received nonzero SNAP benefits in a particular quarter.

- To generate the measure of simulated eligibility that forms one of my dependent variables, I use federal poverty thresholds for each year to determine the SNAP income cutoff for each state in each year if the multiples of poverty used by each state as its BBCE threshold for eligibility were the same as their 2019 levels. This measure relies on the assumption that, while the generosity of SNAP requirements may have changed over time (resulting in this measure over- or underestimating true SNAP eligibility in past years), the relative generosity of states within a given year has remained constant. If this assumption holds true, any bias in the measure should be state-invariant, and is therefore accounted for by year controls.

Some of my robustness checks involve different ways of accounting for the calendar year. There is evidence (e.g., Dube et al. (2010)) that minimum-wage increases may be endogenous to state and national macroeconomic trends. Moreover, minimum-wage increases, and particularly the largest ones, are not evenly distributed across the time period of study; many of the largest ones occurred in the late 2000s (Figure 2), with the periods afterwards characterized by economic downturn. Controlling for fixed effects for each individual year in the period of study, therefore, may eliminate some identifying variation, possibly obscuring a true effect. For this reason, I run my main specification using a linear term for the year;

in alternate specifications, I instead include year fixed effects, or no year term at all.

Another set of robustness checks involve including and excluding each of a number of control variables used in past literature on the minimum wage, the income distribution, and SNAP: logged population (e.g., Dube (2019)); real median family income, mean family size, unemployment rates, and Census-division-specific year fixed effects (e.g., Reich and West, 2015). I also run these regressions using different numbers of leads and lags for the event dummies.

4 Data

My main data sources are the Survey of Income and Program Participation (SIPP) and the minimum wage data from the Washington Center for Equitable Growth (Vaghul and Zipperer, 2016). I also use US Census Bureau data provided by FRED (U.S. Census Bureau, 2020a) for state-level annual population estimates. These datasets are described below.

4.1 SIPP Data

I use SIPP data from January 1996 to October 2013, which includes monthly data on family-level income and SNAP participation. Using the family weights provided in the dataset, I generate several state-level variables, each for the fraction of families that, given their size, falls below a particular multiple of the federal poverty threshold for the relevant year and family size. The multiples of poverty used, as listed above, are 75%, 100%, 125%, 150%, and 200%. I then aggregate each variable to the state-quarter level by generating weighted means over the three months in each quarter.

I also generate similar variables for the fraction of families estimated to be eligible for SNAP, and that of families participating in it. The formula for simulated SNAP eligibility is as described in the empirical strategy section above.

For the first month of my period of study, January 1996, there are a total of 20,142 families surveyed. In the final month, October 2013, there are 14,181 families. Descriptive statistics for these families (as well as families of four, of which there are 2,696 for January 1996 and 1,679 for October 2013) are shown below (all dollar amounts are given in 2019 dollars). In general, the real income distribution has shifted slightly upward, although more so for the upper end of the distribution; and rates of each multiple of poverty have slightly declined. SNAP rates and benefits, however, have increased, likely due to the increased use of BBCE thresholds by states.

	January 1996	October 2013
Median income, all families	\$3,529	\$3,670
25th pctile income, families of four	\$3,290	\$3,554
Median income, families of four	\$6,024	\$6,627
75th pctile income, families of four	\$9,396	\$10,470
% below poverty, families of four	15.3%	14.3%
\$ below 1.5x poverty, families of four	24.8%	23.2%
% below 2x poverty, families of four	33.9%	32.0%
% families of four receiving SNAP benefits	7.11%	11.12%
% Median SNAP benefit per family of four	\$355	\$377

4.2 Minimum Wage Data

This dataset includes the mean, minimum, and maximum values of the federal and state minimum wage in each quarter for every US state and DC from 1996 to 2013. From this I generate a minimum wage change for each state and quarter, defined as the binding minimum wage for the current quarter (the greater of the federal and state minimum wages) minus that for the previous quarter in the same state.

During my time period (1996 to 2013), there were 45 minimum-wage changes, all increases, each applying to anywhere from 1 to 46 states; or a total of 325 state-quarters in which a minimum-wage change occurred. Changes are distributed roughly evenly over all four quarters of the year, meaning that time-from-event is not strongly correlated with the quarter of the year; this allows me to control for seasonal effects without eliminating valuable identifying variation. The mean increase, weighted by number of states covered, was about 8.3%; the middle 50% of affected state-quarters saw increases between 4.1% and 10.5%.

Some of my regression specifications involve only considering minimum wage increases of at least a certain size. A table of these change sizes and their frequencies is shown below:

Size	State-quarters w/ event	Quarters w/ 1+ state event
All changes	325	45
5+ percent	231	35
7.5+ percent	204	34
10+ percent	117	22

Figure 1 shows the number of events of each size in the data. Figure 2 shows a scatter plot of each change in the data; note that most of the largest changes occurred between 2005 and 2010.

5 Results

Results from the main regression specifications (with a linear year term and logged population term) show negative (that is, poverty-reducing) effects of being after a minimum-wage increase, which are significant at the 0.05 alpha level. In general, these effects tend to begin at the 2nd quarter after an increase, grow until the 4th quarter after, and then fade. These results are generally robust to the use of different numbers of pre-/post-event leads and lags.

In the 5-percent version of the regressions (Tables 3a-3b), the effects are significant and negative for the second, third, and fourth quarters post-event when considering the fractions of families below 75% (Table 3a, column 1), 100% (Table 3a, column 2), 125% (Table 3a, column 3), and 150% (Table 3b, column 1) of the poverty line. The largest effects are seen for the 125% and 150% variables, where we see coefficients of around -0.03. That is, the rates of 125% and 150% poverty are about three percentage points lower beginning one to two quarters after an event and continuing until about a year out. There are no significant post-event effects for 200% of poverty (Table 3b, column 2), or for simulated SNAP eligibility (Table 3b, column 3; threshold ranges from 130% to 200% poverty) or participation (Table 3b, column 4). Only a few pre-event coefficients are significant, suggesting that endogeneity (after accounting for controls) is minimal.

The 7.5-percent version (Tables 4a-4b) yield similar results, with only very slightly larger effects. The 10-percent version (Tables 5a-5b), however, yields no results for the 2nd through 4th quarters post-event, and significant effects in the 7th or 8th quarters for some dependent variables (some are positive and some negative). This may be a result of the limited number of 10-percent-or-greater minimum-wage increases in the period of study.

In the version (6a-6b) where minimum-wage increases are weighted by the percent increase they represent, the same pattern is apparent: significant negative effects in the 2nd to 4th quarters post-event, for the 75%, 100%, 125%, and 150% variables only. Again, the largest effects were found in the 125% and 150% columns (Table 6a, column 3; Table 6b, column 1), with coefficients around -0.004. That is, increasing the size of a minimum-wage increase (as a percent of the existing minimum wage) by a single percentage point – for example, going from $12.00 to $12.12 – reduces the rates of 125% and 150% poverty (within the 2nd-to-4th-quarter period) by about 0.4 percentage points.

5.1 Robustness Checks

The above results are generally robust to changing the number of leads and lags; the inclusion and exclusion of each control variable; as well as different ways of accounting for yearly variation (controlling for year fixed effects and leaving out year controls entirely). However, in the fixed-effects version, the magnitudes of the coefficients are generally smaller (sometimes half or less as large as the corresponding ones in the main tables); and some coefficients are not significant. This is consistent with the idea that year fixed effects may remove some of the identifying variation. Moreover, in the fixed-effects version, the results are sensitive to the number of leads and lags used.

5.2 Comparison to Reich and West (2015)

The results here contradict the findings of Reich and West (2015), who use an annual state panel approach to study the effect of the minimum wage on SNAP participation, and find that a 10% increase in the minimum wage is associated with a 2.4–3.2% reduction in SNAP participation.

To investigate why this may have occurred, I replicate their strategy using the state-quarter-level data used in my event-study approach. I run the following specification:

$$Y_{s,t} = \beta_0 + \beta_1 \log(RealMinimumWage_{s,t}) + \beta_2 Controls_{s,t} + \alpha_s + \gamma_q + \phi_{d,y} + \epsilon_{s,t}$$

where $Y_{s,t}$ is the SNAP participation rate in state s and time t; $Controls_{s,t}$ closely approximates the vector of controls used by Reich and West in their study; α_s and γ_q represent state and quarter-within-year fixed effects as before; and $\phi_{d,y}$ represents year fixed effects that vary by Census division.

The result of this regression is significant and negative, indicating that higher minimum wages are associated with lower SNAP participation. This result (Table 7) holds whether the regression is run on the full quarterly panel (columns 1-2) or an annualized version (columns 3-4); and is robust to the use of either linear or fixed-effect Census-division-specific year controls. Coefficients range from -0.024 to -0.030; these are similar to those found by Reich and West (2015), who obtain a coefficient of -0.031. These results imply that increasing the minimum wage by 10% increase reduces SNAP participation by about 0.3 percentage points – that is, about 3% if the SNAP participation is 10%.

This suggests that my findings contradict those of Reich and West (2015), not due to differences in data or time period, but because of the different empirical strategies used.

6 Discussion, Limitations, and Future Work

Overall, the results indicate that increasing the minimum wage increases income for families below 150% of the poverty line, and particular those below that line but above the poverty line itself; and that this effect persists for about a year after the wage increase. There is some indication that smaller wage increases (around 5%) are about as effective as larger ones (around 7.5% – half again as large); however, there is insufficient evidence to conclude that this holds true across the spectrum of possible wage-hike sizes. There is no evidence that larger or smaller minimum wage increases affect different portions of the income distribution.

There is consistently no evidence that the minimum wage reduces either simulated SNAP eligibility or participation at any time. This, combined with the fact that SNAP cutoffs fall between 130% and 200% of the poverty line, suggests that families just below the threshold for SNAP eligibility may be above the income range that benefits most from minimum wage increases; and that these increases, although they boost income, do not do so to the point of lifting families out of SNAP eligibility. This contradicts the finding of Reich and West (2015) that higher minimum wages are associated with lower SNAP participation rates. Moreover, replicating their approach with the data used for this study reproduces their finding, indicating that the difference in results arises from differences between the state panel and event study approaches.

The delayed beginning of the effects is counterintuitive, since the theoretical prediction is that the increased wages would boost families' income immediately, and then the effects would be reduced as the labor market responded to the higher costs of labor. The dissipation of the effect over time, meanwhile, could be driven by a number of factors. It could be that, as inflation reduces the value of the minimum wage and eventually makes it equivalent in real terms to its pre-increase value, the effect of the increase goes away. Firms may also take time to substitute toward capital or non-minimum-wage labor – in this case, some of the income-reducing labor demand effects of the minimum wage would take a while to take effect and offset the increase in income that initially comes from a higher wage. Finally, the dissipation may be partially accounted for by the findings of Neumark et al. (2004), who find that after the first year following a minimum-wage increase, employers are less likely to increase wages, and earned income tends to decrease.

One limitation of this study is in the potential limitations of the SIPP data used to construct state-level variables. Using the family weights provided, the SIPP is capable of producing estimates that are representative at the state level for some but not all states from the year 2004 on (U.S. Census Bureau, 2020b). However, this leaves room for flawed estimates of the rates of low income and SNAP participation in some states and in the earlier years of the time period I consider. If the errors potentially introduced by this are uncorrelated with changes in the minimum wage, then they simply add random noise to my estimates, leading to greater standard errors. If, however, the errors are systematically correlated with minimum-wage changes, they may have also introduced bias in my results. Future work on this topic might rerun the same regressions using data from a source designed to be representative at the state level.

Another limitation is that the measure of simulated SNAP eligibility used is a relatively simplistic one, which does not take into account the possibility of differing state-level trends in SNAP income cutoffs over time, or other state-level criteria beyond income that affect eligibility. Thus this study's results on simulated SNAP eligibility provide only limited information on the effects of the minimum wage on true SNAP eligibility as well as takeup rates. Future research in this area might collect more detailed data on historical state-level SNAP eligibility policies, and use it to construct a more sophisticated measure of simulated eligibility.

Additional possible directions for future research might focus on the mechanisms underlying the dissipation of the effect over time; or use the same event-study strategy to study participation in another social safety net program. Future methodological work might also further examine the differences between the state panel and event study approaches in the context of minimum wage research, and the potential biases to which each approach is subject.

References

Adams, C., Meer, J., and Sloan, C. (2018). The minimum wage and search effort. NBER Working Paper no. 25128.

Addison, J. T. and Blackburn, M. (1999). Minimum wages and poverty. *ILR Review*, 52(3):393–409.

Allegretto, S., Dube, A., Reich, M., and Zipperer, B. (2017). Credible research designs for minimum wage studies: A response to Neumark, Salas, and Wascher. *ILR Review*, 70(3):559–592.

Belman, D., Wolfson, P., and Nawakitphaitoon, K. (2015). Who is affected by the minimum wage? *Industrial Relations: A Journal of Economy and Society*, 54(4):582–621.

Card, D. (1992). Using regional variation in wages to measure the effects of the federal minimum wage. *ILR Review*, 46(1):22–37.

Card, D. and Krueger, A. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4):772–793.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs: Evidence from the United States using a bunching estimator. NBER Working Paper no. 25434.

Clemens, J. and Strain, M. R. (2017). *Estimating the employment effects of recent minimum wage changes: Early evidence, an interpretative framework, and a pre-commitment to future analysis.* AEI Paper & Studies.

Dube, A. (2019). Minimum wages and the distribution of family incomes. *American Economic Journal: Applied Economics*, 11(4):268–304.

Dube, A., Lester, T. W., and Reich, M. (2010). Minimum wage effects across state borders. *The Review of Economics and Statistics*, 92(4):945–964.

Fields, G. S. and Kanbur, R. (2007). Minimum wages and poverty with income-sharing. *The Journal of Economic Inequality*, 5(2):135–147.

Leigh, A. (2007). Does raising the minimum wage help the poor? *Economic Record*, 83(263):432–445.

Meer, J. and West, J. (2016). Effects of the minimum wage on employment dynamics. *Journal of Human Resources*, 51(2):500–522.

Neumark, D., Schweitzer, M., and Wascher, W. (2004). Minimum wage effects throughout the wage distribution. *Journal of Human Resources*, 39(2):425–450.

Neumark, D., Schweitzer, M., and Wascher, W. (2005). The effects of minimum wages on the distribution of family incomes: A non-parametric analysis. *Journal of Human Resources*, 40(4):867–894.

Neumark, D. and Wascher, W. (1992). Employment effects of minimum and subminimum wages: Panel data on state minimum wage laws. *ILR Review*, 46(1):55–81.

Neumark, D. and Wascher, W. (2002). Do minimum wages fight poverty? *Economic Inquiry*, 40(3):315–333.

Phelan, B. J. (2019). Hedonic-based labor supply substitution and the ripple effect of minimum wages. *Journal of Labor Economics*, 37(3):905–947.

Reich, M. and West, R. (2015). The effects of minimum wages on food stamp enrollment and expenditures. *Industrial Relations*, 54(4):668–694.

Sabia, J. (2008). Minimum wages and the economic well-being of single mothers. *Journal of Policy Analysis and Management*, 27(4):848–866.

Sabia, J. (2009). The effects of minimum wage increases on retail employment and hours: New evidence from monthly CPS data. *Journal of Labor Research*, 30(1):75–97.

Sandler, D. H. and Sandler, R. (2014). Multiple event studies in public finance and labor economics: A simulation study with applications. *Journal of Economic and Social Measurement*, 39:31–57.

U.S. Census Bureau (2020a). Annual state population estimates. Retrieved from FRED, Federal Reserve Bank of St. Louis, February 3, 2020.

U.S. Census Bureau (2020b). *Surveys and Programs.* Retrieved from https://www.census.gov/topics/income-poverty/poverty/guidance/surveys-programs.html.

Vaghul, K. and Zipperer, B. (2016). *Historical state and sub-state minimum wage data.* Washington Center for Equitable Growth.

Vedder, R. and Gallaway, L. (2002). The minimum wage and poverty among full-time workers. *Journal of Labor Research*, 23(1):41–49.

Watson, T. (2019). SNAP benefits and the government shutdown.

Wolfson, P. and Belman, D. (2004). The minimum wage: Consequences for prices and quantities in low-wage labor markets. *Journal of Business & Economic Statistics*, 22(3):296–311.

Tables

# of people	SNAP cutoff	Equivalent wage for full-time worker
1	$20,040/year	$9.63/hour
2	$27,168/year	$13.06/hour (one worker) $6.53/hour (two workers)
3	$35,196/year	$16.92/hour (one worker) $8.46/hour (two workers)
4	$41,424/year	$9.96/hour (two workers)

Table 1: 2019 SNAP gross income limits in the state of Minnesota (set at 165% of the federal poverty line), and the hourly wage at which one or two full-time workers would earn exactly that limit. Minnesota's minimum wage was $9.86 in 2019, implying that the minimum wage would keep some families but not others out of SNAP eligibility. Source: Minnesota Department of Human Services (2019).

SNAP income limit set by broad-based categorical eligibility (BBCE) (% of poverty line)	States
130%/no BBCE	Alabama, Alaska, Arkansas, Georgia, Idaho*, Indiana*, Kansas, Kentucky, Louisiana, Mississippi, Missouri, Nebraska*, Ohio, Oklahoma, South Carolina, South Dakota, Tennessee, Utah, Virginia, Wyoming
160%	Iowa, Pennsylvania
165%	Illinois, Minnesota, New Mexico
185%	Arizona, Connecticut, Maine*, New Hampshire, New Jersey, Oregon, Rhode Island, Texas*, Vermont
200%	California, Colorado, Delaware, Florida, Hawaii, Maryland, Massachusetts, Michigan*, Montana, Nevada, New York, North Carolina, North Dakota, Washington, West Virginia, Wisconsin

Table 2: BBCE income cutoffs in each state, as percentages of the poverty threshold below which a family in that state categorically qualifies for SNAP. In states with a *, the higher BBCE income limit only applies to those who fall below a certain asset limit (varies by state). Additionally, the higher limits in New Hampshire and New York only apply to families with dependents. Source: USDA FNS (2019).

VARIABLES	(1) 75% of Poverty	(2) 100% of Poverty	(3) 125% of Poverty
9 before	-0.00542**	-0.00785***	-0.00715**
	(0.00226)	(0.00259)	(0.00344)
8 before	-0.00161	-0.00356	-0.00628
	(0.00452)	(0.00472)	(0.00473)
7 before	-0.000476	-0.00238	-0.00303
	(0.00271)	(0.00381)	(0.00518)
6 before	-0.00245	-0.00437	-0.00765
	(0.00318)	(0.00477)	(0.00674)
5 before	0.00391	0.00518*	0.00314
	(0.00309)	(0.00304)	(0.00373)
4 before	-0.0110	-0.00735	-0.00504
	(0.0111)	(0.00870)	(0.00702)
3 before	-0.0111	-0.0127*	-0.0113
	(0.00697)	(0.00667)	(0.00998)
2 before	-0.00447	-0.000846	0.00111
	(0.00365)	(0.00579)	(0.00865)
1 before	-0.0153*	-0.00959*	-0.00661
	(0.00777)	(0.00486)	(0.00750)
Event time	-0.00191	-0.00440	-0.00285
	(0.00499)	(0.00514)	(0.00866)
1 after	-0.0121*	-0.00635	-0.00498
	(0.00616)	(0.00597)	(0.00481)
2 after	-0.0197***	-0.0228***	-0.0257***
	(0.00595)	(0.00573)	(0.00690)
3 after	-0.0194***	-0.0251***	-0.0304***
	(0.00585)	(0.00613)	(0.0105)
4 after	-0.0104***	-0.00665	-0.0174**
	(0.00306)	(0.00457)	(0.00758)
5 after	0.00582	0.00611	0.00445
	(0.00673)	(0.00731)	(0.00789)
6 after	-0.00369	-0.00301	-0.00728
	(0.00995)	(0.00760)	(0.00680)
7 after	-0.00481	-0.00541	-0.00620
	(0.00657)	(0.00667)	(0.00779)
8 after	-0.0154	-0.0131	-0.0135
	(0.0121)	(0.0107)	(0.0120)
9 after	-0.00841	-0.00947	-0.0107
	(0.00672)	(0.00660)	(0.00741)
Observations	543	543	543
R-squared	0.273	0.306	0.318
Number of statefips	49	49	49

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3a: Results of event-study regressions of the fractions of families that fall below 75%, 100%, and 125% of the poverty line. Events are defined as minimum-wage increases of at least 5%. Significant negative effects for the 2nd through 4th quarters ($p < 0.05$) for all dependent variables.

VARIABLES	(1) 150% of Poverty	(2) 200% of Poverty	(3) SNAP Elig.	(4) SNAP Partic.
9 before	-0.00579*	-0.00508	-0.00658*	-0.00292
	(0.00344)	(0.00432)	(0.00372)	(0.00188)
8 before	-0.00670	-0.0117	-0.0108	-0.00148
	(0.00652)	(0.00713)	(0.00657)	(0.00231)
7 before	-0.00667	-0.0113	-0.0113	0.00167
	(0.00680)	(0.00736)	(0.00675)	(0.00266)
6 before	-0.00988	-0.0125	-0.0161*	-0.00199
	(0.00750)	(0.00948)	(0.00920)	(0.00288)
5 before	-0.000276	0.000177	-0.00252	-0.00116
	(0.00459)	(0.00828)	(0.00869)	(0.00344)
4 before	-0.00743	-0.00902	-0.00952	-0.00148
	(0.00852)	(0.00982)	(0.00956)	(0.00242)
3 before	-0.00798	-0.0137	-0.0114	-0.00506
	(0.00808)	(0.0100)	(0.00926)	(0.00466)
2 before	0.00199	-0.00173	-4.97e-05	-9.98e-05
	(0.00685)	(0.00792)	(0.00762)	(0.00255)
1 before	-0.00575	-0.00348	-0.000165	0.000999
	(0.00720)	(0.00631)	(0.00676)	(0.00218)
Event time	-0.00119	-0.00160	-0.000122	0.00261
	(0.00978)	(0.0126)	(0.0120)	(0.00248)
1 after	-0.00705	0.00348	0.00179	0.00555
	(0.00898)	(0.0123)	(0.0118)	(0.00379)
2 after	-0.0300**	-0.00349	-0.00780	0.00375
	(0.0131)	(0.0156)	(0.0153)	(0.00316)
3 after	-0.0321**	-0.0111	-0.0205	-0.00505
	(0.0122)	(0.0201)	(0.0213)	(0.00600)
4 after	-0.0219***	-0.00954	-0.0111	-0.0107**
	(0.00793)	(0.00978)	(0.00689)	(0.00503)
5 after	0.0111	0.0120	0.0130	-0.00157
	(0.00965)	(0.0108)	(0.0101)	(0.00747)
6 after	-0.00423	0.00279	0.00296	0.00393
	(0.00754)	(0.0109)	(0.00910)	(0.00831)
7 after	-0.00756	-0.00694	-0.00630	-0.00612
	(0.00767)	(0.00961)	(0.00805)	(0.00607)
8 after	-0.0162	-0.0177	-0.0166	-0.00335
	(0.0115)	(0.0124)	(0.0117)	(0.00773)
9 after	-0.0171*	-0.0165	-0.0179*	-0.00658
	(0.00931)	(0.0106)	(0.0102)	(0.00722)
Observations	543	543	543	543
R-squared	0.345	0.262	0.295	0.731
Number of statefips	49	49	49	49

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3b: Results of event-study regressions of the fractions of families that fall below 150% and 200% of the poverty line, as well as simulated SNAP eligibility rate and SNAP participation rate. Events are defined as minimum-wage increases of at least 5%. Significant negative effects for the 2nd through 4th quarters ($p < 0.05$) for 150% of poverty only; and for SNAP participation in the 4th quarter.

VARIABLES	(1) 75% of Poverty	(2) 100% of Poverty	(3) 125% of Poverty
9 before	-0.00391	-0.00599**	-0.00519
	(0.00275)	(0.00297)	(0.00372)
8 before	0.00196	-0.00140	-0.00322
	(0.00469)	(0.00516)	(0.00480)
7 before	0.00102	-0.00133	-0.000334
	(0.00315)	(0.00423)	(0.00564)
6 before	-0.000242	-0.000460	-0.00208
	(0.00342)	(0.00486)	(0.00707)
5 before	0.00533	0.00836**	0.00725*
	(0.00353)	(0.00342)	(0.00406)
4 before	0.00354	0.00417	0.00614
	(0.00338)	(0.00281)	(0.00484)
3 before	-0.00243	-0.00451	-0.00696
	(0.00315)	(0.00368)	(0.00753)
2 before	-0.00670*	-0.00660*	-0.00763
	(0.00353)	(0.00357)	(0.00593)
1 before	-0.0127*	-0.0129**	-0.0135
	(0.00736)	(0.00552)	(0.00874)
Event time	-0.00397	-0.00330	-0.0106
	(0.00409)	(0.00451)	(0.00790)
1 after	-0.00845**	-7.24e-06	-0.00503
	(0.00414)	(0.00411)	(0.00570)
2 after	-0.0262***	-0.0276***	-0.0338**
	(0.00484)	(0.00871)	(0.0147)
3 after	-0.0215***	-0.0272***	-0.0347**
	(0.00555)	(0.00401)	(0.0153)
4 after	-0.0209***	-0.0113***	-0.0312***
	(0.00317)	(0.00333)	(0.00588)
5 after	0.00571	0.00712	0.00524
	(0.00993)	(0.0109)	(0.0118)
6 after	0.00319	0.00346	-0.00108
	(0.0101)	(0.00895)	(0.00925)
7 after	0.00307	0.00233	0.000432
	(0.00899)	(0.00878)	(0.00957)
8 after	-0.000256	-0.0129*	0.00442
	(0.00710)	(0.00679)	(0.00768)
9 after	-0.00302	-0.00588	-0.00793
	(0.00786)	(0.00843)	(0.00812)
Observations	543	543	543
R-squared	0.255	0.292	0.309
Number of statefips	49	49	49

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4a: Results of event-study regressions of the fractions of families that fall below 75%, 100%, and 125% of the poverty line. Events are defined as minimum-wage increases of at least 7.5%. Significant negative effects for the 2nd through 4th quarters ($p < 0.05$) for all dependent variables.

VARIABLES	(1) 150% of Poverty	(2) 200% of Poverty	(3) SNAP Elig.	(4) SNAP Partic.
9 before	-0.00390	-0.00527	-0.00709*	-0.00286
	(0.00390)	(0.00459)	(0.00396)	(0.00209)
8 before	-0.00297	-0.00711	-0.00647	-0.00117
	(0.00685)	(0.00718)	(0.00660)	(0.00257)
7 before	-0.00289	-0.00882	-0.00875	0.000958
	(0.00734)	(0.00762)	(0.00702)	(0.00256)
6 before	-0.00478	-0.0116	-0.0153*	-0.00224
	(0.00727)	(0.00947)	(0.00903)	(0.00326)
5 before	0.00519	0.000743	-0.00114	-0.00292
	(0.00473)	(0.00565)	(0.00522)	(0.00376)
4 before	0.00601	0.00281	0.00175	-4.05e-05
	(0.00550)	(0.00792)	(0.00781)	(0.00261)
3 before	-0.00444	-0.00808	-0.00647	-0.00444
	(0.00574)	(0.00807)	(0.00754)	(0.00441)
2 before	-0.00429	-0.00177	-0.00191	-0.000216
	(0.00604)	(0.00825)	(0.00823)	(0.00318)
1 before	-0.0111	-0.00187	-0.00115	0.00125
	(0.00821)	(0.00587)	(0.00567)	(0.00234)
Event time	-0.0126	-0.0133	-0.0131	-0.000273
	(0.00769)	(0.0118)	(0.00903)	(0.00225)
1 after	-0.00652	0.00302	-0.000644	0.00550
	(0.0149)	(0.0203)	(0.0196)	(0.00599)
2 after	-0.0331	0.0102	0.00282	0.00185
	(0.0289)	(0.0283)	(0.0303)	(0.00734)
3 after	-0.0398*	0.000664	-0.0111	0.000929
	(0.0209)	(0.0343)	(0.0387)	(0.00502)
4 after	-0.0399***	-0.0248**	-0.0199*	-0.00916***
	(0.00797)	(0.00982)	(0.0100)	(0.00305)
5 after	0.0105	0.0119	0.0140	-0.00421
	(0.0110)	(0.0118)	(0.0106)	(0.00628)
6 after	0.000743	0.00413	0.00583	0.00284
	(0.00905)	(0.0119)	(0.0104)	(0.00726)
7 after	0.00189	-0.00209	0.000522	-0.00195
	(0.0102)	(0.0120)	(0.00985)	(0.00686)
8 after	0.00759	-0.0322***	-0.0167**	0.0174***
	(0.00689)	(0.00818)	(0.00752)	(0.00557)
9 after	-0.0102	-0.0125	-0.0141**	0.000632
	(0.00767)	(0.00796)	(0.00672)	(0.00617)
Observations	543	543	543	543
R-squared	0.331	0.253	0.282	0.727
Number of statefips	49	49	49	49

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4b: Results of event-study regressions of the fractions of families that fall below 150% and 200% of the poverty line, as well as simulated SNAP eligibility rate and SNAP participation rate. Events are defined as minimum-wage increases of at least 7.5%. Significant negative effects in the 4th quarter ($p < 0.05$) for all dependent variables; and in the 3rd quarter for 150% of poverty.

VARIABLES	(1) 75% of Poverty	(2) 100% of Poverty	(3) 125% of Poverty
9 before	-0.00398	-0.00561*	-0.00530
	(0.00254)	(0.00287)	(0.00365)
8 before	0.000730	-0.00312	-0.00324
	(0.00436)	(0.00481)	(0.00485)
7 before	0.000162	-0.00284	-0.000123
	(0.00277)	(0.00384)	(0.00576)
6 before	0.000259	-0.000313	-3.97e-05
	(0.00315)	(0.00470)	(0.00727)
5 before	0.00423	0.00703**	0.00809**
	(0.00268)	(0.00284)	(0.00394)
4 before	0.00380	0.00471	0.00690
	(0.00350)	(0.00307)	(0.00531)
3 before	-0.00197	-0.00174	0.000583
	(0.00302)	(0.00296)	(0.00507)
2 before	-0.00723**	-0.00702*	-0.00468
	(0.00348)	(0.00371)	(0.00577)
1 before	-0.00656	-0.00877***	-0.00470
	(0.00443)	(0.00327)	(0.00399)
Event time	-0.00283	-0.00455	-0.00187
	(0.00438)	(0.00506)	(0.00366)
1 after	-0.00260	-0.00545	-0.000424
	(0.00292)	(0.00460)	(0.00247)
7 after	0.0318***	0.0202***	0.0204***
	(0.00577)	(0.00544)	(0.00578)
8 after	0.00407	-0.0102**	0.00767
	(0.00519)	(0.00489)	(0.00526)
9 after	0.00428	0.000851	-0.00288
	(0.00524)	(0.00623)	(0.00682)
Observations	543	543	543
R-squared	0.245	0.280	0.295
Number of statefips	49	49	49

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5a: Results of event-study regressions of the fractions of families that fall below 75%, 100%, and 125% of the poverty line. Events are defined as minimum-wage increases of at least 10%. Significant negative effects in the 7th quarter ($p < 0.05$) for all dependent variables. Some significant pre-event coefficients, casting doubt on results.

VARIABLES	(1) 150% of Poverty	(2) 200% of Poverty	(3) SNAP Elig.	(4) SNAP Partic.
9 before	-0.00394	-0.00572	-0.00774**	-0.00220
	(0.00380)	(0.00460)	(0.00383)	(0.00195)
8 before	-0.00141	-0.00836	-0.00783	-0.00131
	(0.00727)	(0.00756)	(0.00694)	(0.00265)
7 before	-0.00151	-0.00964	-0.00969	0.000169
	(0.00770)	(0.00848)	(0.00786)	(0.00261)
6 before	-0.00244	-0.0107	-0.0145	-0.00274
	(0.00770)	(0.0107)	(0.0101)	(0.00344)
5 before	0.00620	0.00189	0.000199	-0.00278
	(0.00478)	(0.00608)	(0.00577)	(0.00391)
4 before	0.00624	0.00162	0.000623	0.000768
	(0.00607)	(0.00756)	(0.00777)	(0.00273)
3 before	0.00136	-0.000602	0.000117	0.000750
	(0.00473)	(0.00674)	(0.00624)	(0.00248)
2 before	-0.00345	-0.00229	-0.00353	0.00253
	(0.00605)	(0.00857)	(0.00848)	(0.00181)
1 before	-0.00342	-0.00219	-0.00431	0.00139
	(0.00413)	(0.00629)	(0.00601)	(0.00204)
Event time	-0.00442	-0.00252	-0.00511	-0.00141
	(0.00388)	(0.00555)	(0.00381)	(0.00162)
1 after	-0.00484*	-0.000486	-0.00390	-0.00412
	(0.00288)	(0.00562)	(0.00435)	(0.00307)
7 after	0.0121**	-0.00855	-0.000930	-0.000890
	(0.00498)	(0.00653)	(0.00577)	(0.00395)
8 after	0.00922**	-0.0333***	-0.0178***	0.0177***
	(0.00415)	(0.00536)	(0.00455)	(0.00399)
9 after	-0.00401	-0.00600	-0.00655	-0.00273
	(0.00705)	(0.00826)	(0.00711)	(0.00600)
Observations	543	543	543	543
R-squared	0.315	0.249	0.278	0.725
Number of statefips	49	49	49	49

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5b: Results of event-study regressions of the fractions of families that fall below 150% and 200% of the poverty line, as well as simulated SNAP eligibility rate and SNAP participation rate. Events are defined as minimum-wage increases of at least 10%. Significant negative effects in the 7th and 8th quarters ($p < 0.05$) for all dependent variables. Some variables were omitted by Stata.

VARIABLES	(1) 75% of Poverty	(2) 100% of Poverty	(3) 125% of Poverty
9 before	-0.000322**	-0.000450**	-0.000392*
	(0.000148)	(0.000178)	(0.000225)
8 before	-0.000130	-0.000282	-0.000266
	(0.000217)	(0.000256)	(0.000254)
7 before	-0.000170	-0.000310	-0.000176
	(0.000143)	(0.000203)	(0.000284)
6 before	-0.000208	-0.000293	-0.000356
	(0.000235)	(0.000309)	(0.000447)
5 before	2.18e-06	0.000197	0.000198
	(0.000389)	(0.000358)	(0.000349)
4 before	-0.000352	-0.000111	-4.38e-05
	(0.000555)	(0.000442)	(0.000424)
3 before	-0.000560	-0.000574	-0.000552
	(0.000447)	(0.000419)	(0.000551)
2 before	-0.000501*	-0.000336	-0.000258
	(0.000277)	(0.000352)	(0.000464)
1 before	-0.000940*	-0.000731*	-0.000468
	(0.000552)	(0.000391)	(0.000482)
Event time	-0.000328	-0.000399	-0.000257
	(0.000330)	(0.000358)	(0.000427)
1 after	-0.00118*	-0.000647	-0.000481
	(0.000660)	(0.000616)	(0.000450)
2 after	-0.00266***	-0.00285***	-0.00336***
	(0.000583)	(0.000709)	(0.000944)
3 after	-0.00297***	-0.00345***	-0.00425***
	(0.000766)	(0.000600)	(0.00103)
4 after	-0.00208***	-0.00141**	-0.00231**
	(0.000497)	(0.000676)	(0.000883)
5 after	-0.000116	4.76e-05	-0.000214
	(0.000617)	(0.000755)	(0.000866)
6 after	-0.000710	-0.000591	-0.00111
	(0.00102)	(0.000815)	(0.000788)
7 after	-0.000399	-0.000435	-0.000598
	(0.000783)	(0.000810)	(0.000916)
8 after	-0.00113	-0.00115	-0.000611
	(0.00114)	(0.00101)	(0.00119)
9 after	-0.000323	-0.000548	-0.000722
	(0.000608)	(0.000644)	(0.000748)
Observations	543	543	543
R-squared	0.265	0.300	0.316
Number of statefips	49	49	49

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6a: Results of event-study regressions of the fractions of families that fall below 75%, 100%, and 125% of the poverty line. All wage-increase events are included and are weighted according to the percentage size of the increase. Significant negative effects for the 2nd through 4th quarters ($p < 0.05$) for all dependent variables.

VARIABLES	(1) 150% of Poverty	(2) 200% of Poverty	(3) SNAP Elig.	(4) SNAP Partic.
9 before	-0.000335	-0.000460	-0.000496*	-0.000157
	(0.000235)	(0.000299)	(0.000263)	(0.000126)
8 before	-0.000232	-0.000539	-0.000500	-0.000106
	(0.000357)	(0.000373)	(0.000328)	(0.000138)
7 before	-0.000318	-0.000655	-0.000629*	-2.42e-05
	(0.000369)	(0.000411)	(0.000367)	(0.000136)
6 before	-0.000494	-0.000782	-0.00104*	-0.000192
	(0.000464)	(0.000590)	(0.000567)	(0.000196)
5 before	5.76e-06	1.92e-05	-0.000205	-0.000143
	(0.000372)	(0.000457)	(0.000449)	(0.000296)
4 before	-0.000235	-0.000392	-0.000429	-0.000112
	(0.000511)	(0.000596)	(0.000606)	(0.000230)
3 before	-0.000502	-0.000735	-0.000541	-0.000354
	(0.000549)	(0.000656)	(0.000625)	(0.000368)
2 before	-0.000166	-0.000354	-0.000317	-2.92e-05
	(0.000436)	(0.000546)	(0.000529)	(0.000183)
1 before	-0.000412	-0.000214	-0.000110	-4.16e-05
	(0.000451)	(0.000473)	(0.000464)	(0.000173)
Event time	-0.000289	-0.000216	-0.000273	2.09e-05
	(0.000478)	(0.000655)	(0.000590)	(0.000190)
1 after	-0.000853	8.18e-05	-0.000157	5.56e-05
	(0.000862)	(0.00107)	(0.00107)	(0.000457)
2 after	-0.00376*	-0.000308	-0.000891	-4.89e-07
	(0.00197)	(0.00202)	(0.00213)	(0.000446)
3 after	-0.00450***	-0.00170	-0.00275	-0.000563
	(0.00143)	(0.00260)	(0.00287)	(0.000717)
4 after	-0.00283**	-0.000882	-0.00106	-0.00138*
	(0.00112)	(0.00137)	(0.00117)	(0.000787)
5 after	0.000606	0.000669	0.000773	-0.000180
	(0.000828)	(0.000964)	(0.000877)	(0.000775)
6 after	-0.000650	-4.85e-06	6.84e-05	0.000294
	(0.000858)	(0.00121)	(0.00104)	(0.000903)
7 after	-0.000797	-0.00130	-0.000996	-0.000667
	(0.000974)	(0.00120)	(0.000992)	(0.000706)
8 after	-0.00103	-0.00236*	-0.00190	-0.000237
	(0.00126)	(0.00131)	(0.00123)	(0.000943)
9 after	-0.00145*	-0.00160	-0.00161*	-0.000683
	(0.000803)	(0.000974)	(0.000906)	(0.000689)
Observations	543	543	543	543
R-squared	0.339	0.260	0.290	0.729
Number of statefips	49	49	49	49

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6b: Results of event-study regressions of the fractions of families that fall below 150% and 200% of the poverty line, as well as simulated SNAP eligibility rate and SNAP participation rate. All wage-increase events are included and are weighted according to the percentage size of the increase. Significant negative effects for the 2nd through 4th quarters ($p < 0.05$) for 150% of poverty.

VARIABLES	(1)	(2)	(3)	(4)
Log minimum wage	-0.0238***	-0.0238***	-0.0301**	-0.0301**
	(0.00661)	(0.00661)	(0.0130)	(0.0130)
Constant	0.203**	0.203**	0.231	0.231
	(0.0944)	(0.0944)	(0.165)	(0.165)
Observations	3,240	3,240	880	880
R-squared	0.697	0.697	0.747	0.747
Number of statefips	48	48	48	48
Year Controls	FE	Linear	FE	Linear
Level	Quarterly	Quarterly	Yearly	Yearly

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7: Results of replication of Reich and West's (2015) state panel study of the effect of the minimum wage on SNAP participation. Columns (1) and (2) are results of quarterly state panels, and columns (3) and (4) are resuls of yearly state panels. Each version is run with both linear and fixed-effect versions of Census-division-specific year controls. Results compare closely to those of Reich and West (2015), who obtain a coefficient of -0.031.

Figures

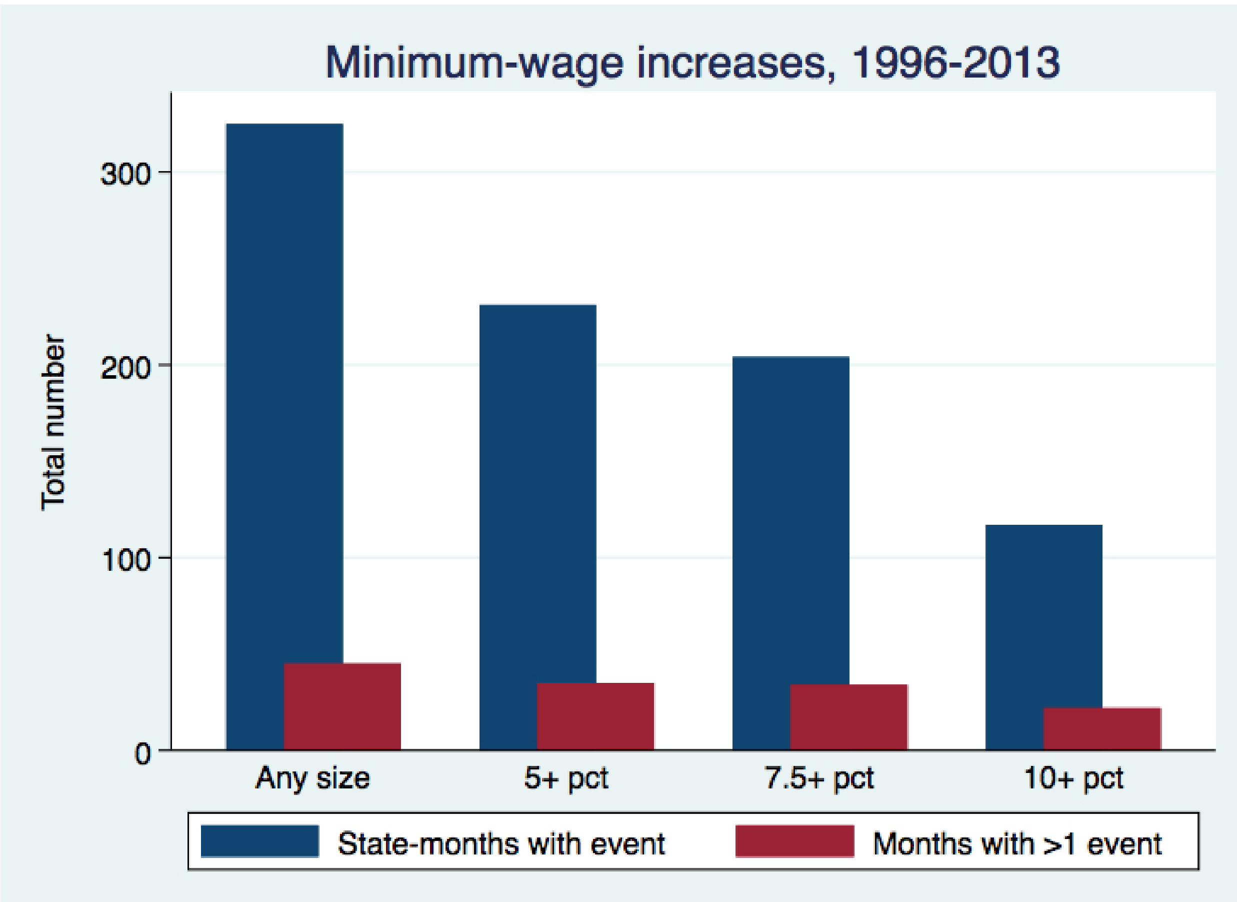


Figure 1: Number of minimum-wage increases of each size throughout the entire dataset (both in terms of state-quarter observations containing increases, and quarters in which at least one state saw an increase).

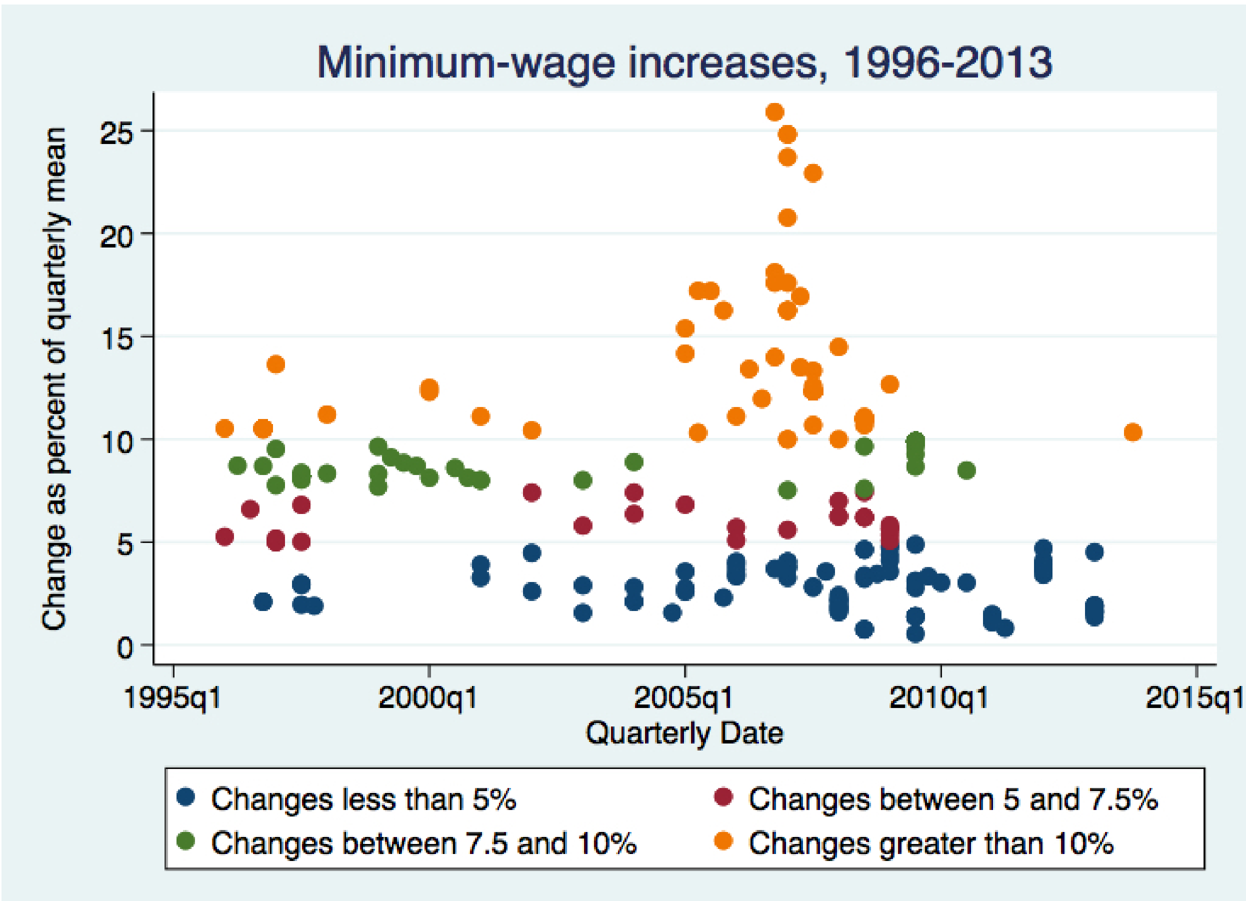


Figure 2: Scatter plot of minimum-wage increases of each size over time.