ESSAYS IN ECONOMICS

by

Chuan (Charles) He

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ECONOMICS

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2015

THE UNIVERSITY OF ARIZONA GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Chuan (Charles) He, titled Essays in Economics and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

______________________________ Date: 06/03/2015
Gautam Gowrisankaran

______________________________ Date: 06/03/2015
Price Fishback

______________________________ Date: 06/03/2015
Ashley Langer

______________________________ Date: 06/03/2015
Mauricio Varela

Final approval and acceptance of this dissertation is contingent upon the candidates submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

______________________________ Date: 06/03/2015
Dissertation Director: Gautam Gowrisankaran

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Chuan (Charles) He

# Contents

# Abstract

This thesis examines various topics of individual choice and welfare in economics. In the first chapter, I examine the choices of workers and business in coal mining. Coal mining is a dangerous occupation where costly fatalities and disasters may increase future accident costs. We use occurrences of deaths as shocks that affect the tradeoff between mineral output and safety. We find that government inspections and penalties increase after fatalities, and less-severe accident rates decrease by 10%. For mines in a disaster-affected state, less-severe accident rates decrease by 23%, and fatalities by 68%, saving up to $2 per hour in accident costs, with limited evidence suggesting that mineral productivity falls by 7%, or $14 per worker hour, and that the number of managers employed increases by 11%.

In the second chapter, I examine how the distribution of welfare from decades of technological improvement in electronics, have benefitted consumers. To do this, I examine the welfare gains to different income cohorts from the development of multiple categories of electronic products. Income dependent preferences are estimated in a dynamic model of demand. Key utility parameters, unique to income cohort, are identified from moments created using micro-level data containing purchase and demographic information. My results suggest that the benefits from the rise of electronics products may be far more egalitarian than conventional measures of inequality. Welfare gains to consumers in the bottom third of the US income distribution average approximately $1,000, while gains to the top third benefit are about $2,500. This is twice as equal as related measures of consumption inequality.

In the last chapter of this dissertation, I examine the influence of early random outcomes on the choices and later success of players in a major online multiplayer game. Players misattribute early random shocks to the value of an important choice.

I find that players are significantly less successful in subsequent matches when making the same choice that benefitted from positive shocks early on. This effect, which occurred in just a handful of initial matches, lasts over the year long career of a player, spanning hundreds of matches of the game. The setting of the game provides a rich dataset, and is also a rare empirical example of a multi-armed bandit problem.

# Chapter 1

# Innovating to Equality: The Egalitarian Distribution of Welfare from the Rise of Consumer Electronics

## 1.1 Introduction

The US consumer electronic industry has grown almost uninterrupted over four decades, and stands at over $200 billion today. The this industry constantly produces new products that contain remarkable innovations and technological accomplishments. However, this industry has another achievement that is far less conspicuous but no less impressive – the widespread use and ownership of many of its products by millions of low income Americans.[1] We have long recognized that improvements in goods and services benefit consumers, and that electronics have been dramatically improved by innovation, a process that shows no signs of slowing down. But far less understood is how the benefits of these recent innovations affect different income groups in our society. How equal are the distribution of the gains? With an eye toward the future, what can we learn in general about the distribution of benefits from these products?

[1] For consumers in households with incomes below $30K, it has been found 73% of these consumers have internet access, 85% have cell phones (35% of which are smart phones) 47% have broadband, and 59% have a desktop or laptop. Source: Pew Internet Project Research.

This paper provides an answer by studying the consumer surplus gains from several major categories of electronics. However, the same rapid pace of change in these markets poses challenges when using methods involving price indexes or consumption data. This suggests the effects of consumer surplus from electronic product development on inequality is both an empirical question, and also motivates its measurement using a demand model that can estimate consumer surplus for these products.

To illustrate the challenges, imagine ways how we might try to observe the benefits from improvements in a given electronics category. We would expect wealthier consumers to benefit more from improvements in higher end products, while lower income groups benefit from improvements in less expensive models, so perhaps we could study how different segments of the market evolve. But the role of products is fluid in the case of consumer electronics; it is common to see a flagship product repurposed as an entry level model the following year. Another strategy may be to compare spending: we would expect gains to be more equal if different income groups spent equal amounts, compared to the case if the wealthy predominantly spent the most. But while there is a relationship between the distribution of spending and the distribution of benefits, plummeting prices can quickly dissociate spending from value: products and features that were premium at the time of their introduction can often be seen a few years later at a tenth of their former price. As a result, low income consumer spending could be small, but the value of their purchases could still be quite high. Alternatively, these products may have been designed to cater to the higher income groups who purchased them original premium prices, and later have much less value to lower income purchasers.

This suggests the need to estimate consumer preferences and willingness to pay over these products. Accordingly, I propose and estimate a model that estimates demand by income cohort, to measure the distribution of welfare to different income groups. There exists only a small literature on welfare from electronics,[2] and to my knowledge, no other papers have examined the inequality of consumer welfare.[3]

Although it may be novel to measure the inequality of welfare, measurements of inequality have been a longstanding topic for economists, e.g. Dalton (1920). There

[2]This literature includes Bresnahan (1986); Hausman et al. (1997); Brynjolfsson (1996); Brynjolfsson et al. (2003). Very recently, Greenwood and Kopecky (2013) estimated surplus from PCs using a simple utility model that relies on the functional form of utility. The age for all but the last paper both demonstrate that economists have long recognized the significance of electronics, and that new estimates are justified.

[3] Note that Petrin (2002) and estimates demand by income group, but only shows and discusses aggregate welfare estimates. His paper informs key parts of the estimation strategy in this paper.

is a large literature examining inequality in income and wealth (e.g. see Juhn et al. (1993); Piketty and Saez (2001); Autor et al. (2008); Meyer and Sullivan (2013)). More recently, motivated by economic theory,[4] a literature has emerged that focuses on consumption inequality (Aguiar and Bils, 2011; Heathcote et al., 2010; Attanasio et al., 2012; Meyer and Sullivan, 2013; Krueger and Perri, 2006). Consumption inequality is an attractive measure of economic inequality because it is more representative of the actual economic variable of interest, consumer welfare itself. Yet despite being fundamental to these literatures, inequality in consumer welfare is rarely ever examined directly. Virtually all studies use variables such as income, consumption, or wealth.

Yet these new products have dramatic rates of improvement that confound conventional measures of consumption measures and related statistics. Conventionally, the value of changing markets is often measured using a cost of living index, which in practice is derived from the CPI. But these indexes struggle to capture the development and improvement in these markets, where new types of product can rise from infancy to widespread adoption in just a few years. This rapid change poses problems. For example, Hausman (1999) criticizes the BLS for ignoring cell phones, whose omission results in considerable biases in the CPI. Yet even perfectly designed, indexes poorly measure the value of new innovations due to inherent limitations. Nordhaus (1996) makes the point "that by the very nature of their construction, price indexes miss the most important technological revolutions in economic history".

The durable nature of these products poses other complications. Durable goods are enjoyed over long periods of time, so consumption is not easily measured by spending. Even if we could somehow perfectly track consumption, other issues emerge. As Melnikov (2013) and Gowrisankaran and Rysman (2012) discuss and show, demand models of consumers durables that do not account for dynamic behavior can perform extremely poorly, because they ignore issues such as intertemporal substitution. Additionally, it is usually the case that consumers already hold some older version of the product, which influences their purchase decisions.

For the reasons above, I use a dynamic structural model that takes into consumer holdings and dynamic behavior such as future expectations. My model follows the line of dynamic frameworks established by Rust (1987) and extended in Melnikov

[4] Consumption better reflects transfers, savings and other forms of insurance than income. See Blundell et al. (2008).

(2013)[5] and Gowrisankaran and Rysman (2012) (abbreviated GR). I describe this briefly below, leaving the remaining detail to the model section.

The core of the estimation is the modelling of consumers as forward looking agents who generally hold existing products from previous purchases. They must decide between purchasing in the current period versus waiting for future improved and less expensive versions of the item. Additionally, my model incorporates two types of heterogeneity. Firstly, I estimate random heterogenous demand in the population using random coefficients, as in Berry et al. (1995) (abbreviated BLP). These random coefficients capture important differences between a broad range of consumers who likely have large variation in their valuations of these products. Secondly, to obtain heterogenous estimates of demand by income, which is the main goal of the paper, I first divide my consumers into three income groups, low, medium and high, that are designed to represent U.S. income terciles. Each group shares a similar utility function, but the parameters of the utility functions vary across income groups. For example, the parameter of price sensitivity varies by income. To estimate these coefficients, the model predicts measures of behavior by group and the estimation uses a process that chooses solutions that fit these predictions to the observed behavior of these consumers.

My model is most similar to the dynamic demand model of GR cited above, and the identification of income heterogeneity is based on Petrin (2002). I provide an identification proof for my implementation of income heterogeneity in the model section. The addition of these features to the model increase the credibility of the estimates of consumer surplus for different income groups in this market for electronic goods.

Two types of data are used by the model. Firstly, I use market level panel sales data of three types of products: camcorders, digital cameras, and DVD players. The panels vary by product, but they span several years from the inception of these products and cover periods of high sales growth and price declines. To identify how preferences among different income groups differ, and to attribute welfare gains to each income group, I use micro-level purchase data from a major retailer that provides joint information about the income of purchasers, and the frequency and prices of their purchases. I use the retailer data to impute the predicted behavior of each income group in my model.

[5] Aside: despite the later publication date, Melnikov is the antecedent of Gowrisankaran and Rysman, who extend his model.

The results show that while surplus is unequal, the distribution is more equal than conventional measures of inequality. Overall, using the average across the three product categories, lower and middle income consumer surplus is 40% and 71%, respectively of the higher income's surplus. The distribution varies by product : camcorders are most equal, digital cameras are the least, and DVD players are in between. As expected, these rankings correspond to the relative share of product purchases. Furthermore, as expected dynamic relationships play an important role in determining the level of surplus. The surplus shows considerable concavity over time; consumers value improvements in the market much less than the sales weighted price changes would suggest, consistent with consumers withholding demand in earlier market states. Also the product with the largest average market improvement, digital cameras, also produces gains that are largest proportional to actual spending, which is consistent with intertemporal substitution and future expectations.

The dynamic demand estimates suggest important relationships between innovation and inequality. I find that digital cameras, the product experiencing the largest average innovation, produces the most unequal surplus. I argue this observation reflects how innovation enters the market of consumers with different economic means. Flows of innovation initially arrive in premium products whose higher prices favor wealthier consumers. As a result, the share of benefits to higher income groups is large in early periods when innovation is high. However, the wealthy's share of benefits decline over time as the flow of innovation falls relative to the past stock of implemented improvements, whose prices fall dramatically. As a result, lower income consumers enjoy a larger proportion of the benefit at more mature stages of the market. Products experiencing high rates of development tend to have a more unequal distribution. Note that one implication of these findings, assuming the rate of innovation is not accelerating, is that if prices decline constantly, in the limit case the gains from innovation will be perfectly egalitarian.

This paper provides estimates of the gains from consumer electronics, and contributes a new perspective to the literature on inequality. The distribution of economic welfare is a topic that many of us have both a personal and academic interest. Some argue that pronounced inequality is not only intrinsically undesirable, but is directly harmful to economic growth and societal welfare.[6]

[6] A rebuttal is made by Becker and Murphy (2007) (AEI opinion article) who argue that inequality reflects variations in human capital and produces desirable incentives. Instead, the focus should be improving the amount and quality of education for children disadvantaged by poor schooling.

Aggregate economic gains may not be sufficient to judge progress because even large growth may prove hollow if a great many do not benefit. Similarly, the importance of the distribution of the benefits from these new products should be understood. These products are both widely and intensively used,[7] and while they may already play significant social and political roles,[8] it is unlikely that their full effects on communications and society are known. Their importance makes the differences between worlds where electronics are used fairly equally, or those where they are mainly enjoyed by the wealthy, seem very stark. By examining its distribution, this paper explores an important but unexamined dimension of economic welfare and inequality. As technologies and markets continue to develop, and the adoption of these products expands, this topic only grows in importance.

The remainder of this paper is organized as follows: subsection 1.1.1 discusses background literature; section 1.2 discusses the model; section 1.3 discusses data and the calculation of welfare; Sections 1.4 and 1.5 contains results and discussion and section 1.6 concludes.

### 1.1.1 Background

Examining the distribution of consumer surplus from new product in these markets improves our understanding of inequality by exploring an otherwise unobservable contributor of economic welfare.[9] To my knowledge, no existing literature examines distributional issues related to consumer welfare, and only by assembling an eclectic literature can one find a small body of work focused on the welfare from electronic goods.

This literature includes Brynjolfsson (1996), who measure welfare improvements from IT spending up to 1987. He finds that by that time, electronics spending

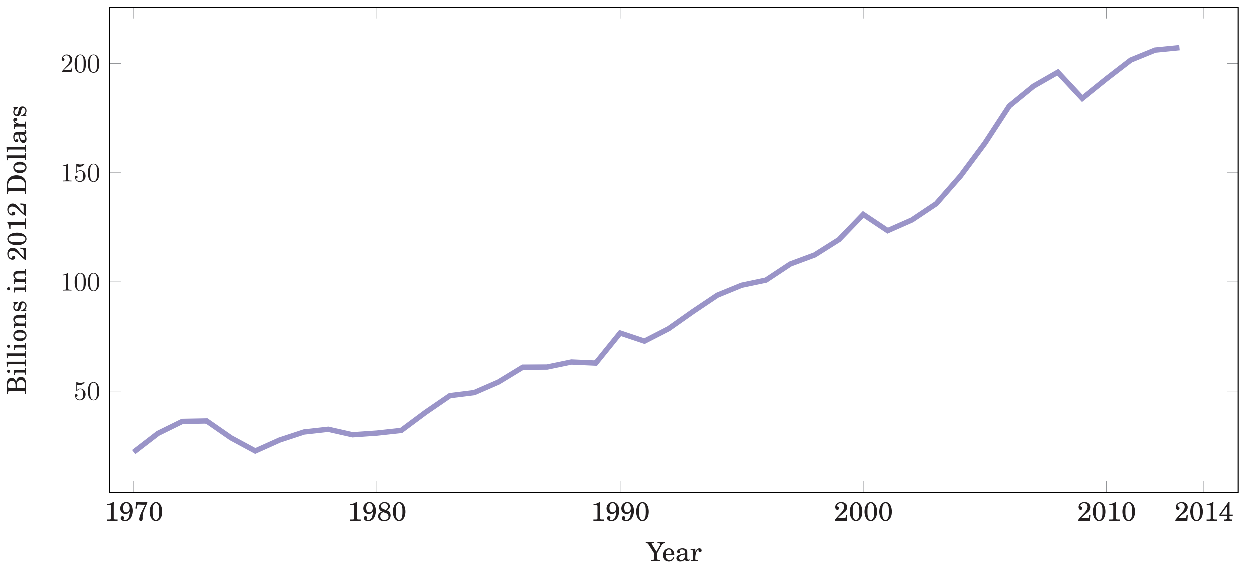
[7] See Fox and Rainie (2014) for statistics and discussions of the rise in computer and internet use and similarly, "Mobile Technology Fact Sheet, 2014", Pew Research Center, for cellular phones.

[8] The societal effects of new technology is so broad and multidisciplinary that I do not attempt a representative summary. However a handful of examples suggests the influence of new communication technology: a role in the Arab spring uprisings (Howard et al., 2011), facilitating secure communications of dissidents Lewman (2013), organizing community responses to disasters (Sutton et al., 2008) and mediating social capital (Ellison et al., 2007). The above does not include potentially far deeper effects that technology might have, for example Fukuyama (2014) argues the printing press created a sense of community identity that was a precursor to the emergence of the nation state.

[9] This is similar in a sense to the recent literature that examines consumption over income as a preferable measure of economic welfare. This literature includes (Krueger and Perri, 2006; Krueger et al., 2010; Aguiar and Bils, 2011; Heathcote et al., 2010; Attanasio et al., 2012; Meyer and Sullivan, 2003, 2008, 2013). This is motivated by economic theory and transfers: consumption reflects taxes, transfers, insurance and other measures of income smoothing (Cutler and Katz, 1992). While measures of inequality using consumption are both theoretically and empirically found to be lower than those measured by income (certainly in levels and possibly in change), in contrast, the relative inequality of welfare effects is unknown.

already generates figures on the order of $100 billion in consumer surplus annually, several times its spending. Recently, Greenwood and Kopecky (2013) measure the benefit from PCs using a different approach and find that PCs are worth 2-3% of the value of consumption, while only accounting for 0.6% of spending. In another electronics category, Brynjolfsson et al. (2003) finds that in 2000, the welfare gain from eBooks was approximately one billion dollars, roughly twice its spending. Finally, an interesting related paper is Goolsbee and Klenow (2006), who estimate the value of the internet by the time spent and find an estimates of up to one thousand of dollars per person, ten times its spending.[10]

Figure 1.1: Annual U.S. Consumer Spending on Electronics




Source: Consumer Electronics Association Market Research (Private Communication).

All papers in this diverse literature find large surpluses relative to spending. Using the lowest estimates, I suggest that a conservative measure of the annual welfare value of gains (in the United States), is gained by setting it equal to spending, shown in Figure 1.1. This produces a figure of $200 billion dollars, or approximately $2000 by household. This is a large figure by many measures; for example, it dwarfs the contemporaneous gain in median household income[11] and is comparable to many

[10] Less directly, some estimates of surplus (which include business purchases) can be obtained from a small literature on microprocessors that find surpluses from these products on the order of one hundred billion annually (Song, 2007; Goettler and Gordon, 2011). Finally, another interesting paper focusing on consumer welfare is by Kuhn and McAusland (2009), who suggest that the marginal contributions of South-Asian engineers emigrants could be large enough that their contributions to consumer welfare in their home countries may reverse losses from brain drain; i.e, South-Asian engineers working in Silicon Valley improve Indian welfare more in than working in India itself.

[11] From 1987 to 2012, median real US income has improved by only half this figure, about $1000, while mean household income has increased by about $10,000.

large government social programs.[12] Compare to the EITC, which distributes about \$2000 per participant or \$200 per capita (Chetty and Saez, 2013). The magnitude of these transfers is large enough to drive various results in studies of school performance and lifetime earnings (Dahl and Lochner, 2008; Marr et al., 2013), teenage pregnancies Chetty and Friedman (2011) and infant health (Hoynes et al., 2012).

While we can get a sense of the magnitude of the gains, there is almost no literature that pins down its distribution. However, there is a sense that some innovation specifically benefits lower income consumers. Generally this is done by providing novel substitutes for services to which these consumers have had historically lower access. One example is mobile phones, a product with rates of adoption that are egalitarian compared to land lines (Goldfarb and Prince, 2008; Brown et al., 2011), and provides another channel for access to the internet. Another example is new banking services, such as Walmart's GoBank, that provide an accessible alternative for under-banked low income consumers (central to these new banking services is the use of mobile and smartphones, as suggested by Gross et al. (2012)). Perhaps the only other paper to identify surplus by income is Petrin (2002), who examines the distribution of welfare from the introduction of the minivan.[13] [14]

There appears to be little theoretical work related to this topic. This is because this problem is likely to be an empirical question, and also perhaps because the mechanism is simple. Setting aside preferences, the relative surplus of different subpopulations is proportional to the expenditures on their subset of purchases, and the relative improvements in the subset of products they purchase. Some informal

[12] Note that this comparison of "innovation" to "income changes" suffers from a form of double counting. This is because innovation affects the price deflator of electronic goods by increasing real income. However, correcting this would make gains from innovation even larger relative to real income gain . Additionally, electronic goods are a small portion of the CPI index: a generous interpretation of the categories the BLS uses to weight the price index in 2013 gives these electronics a five percent weight, which is probably a much greater than in 1987. Therefore the effect on CPI is likely modest, even in the extreme case where the price index of these goods fell to zero.

[13] However, Petrin speaks little to the topic of the paper. The technological change in his paper was limited to a single event and product: the one-time introduction of the minivan. In contrast, innovation in electronics happens continually over many time periods and has unique features that are not captured by a one time event: miniaturization, rapidly declining prices, and the introduction and progressive dissemination of desirable features from high to lower end models.

[14] Other economic literature does recognize the importance of separating analysis by income. In the price index or CPI literature, it is well understood that income groups are affected by price changes differently. This was noted as early as Prais (1959), and examined in the more recent price index literature (Pollak, 1980; Slesnick, 2001; Hobijn and Lagakos, 2005). The literature has proposed adjustments to reflect varying preferences, including stratification (Erbas and Sayers, 1998) and reweighting of purchases to avoid biasing wealthier consumers (Ley, 2005). Howver, the CPI's well known shortcomings (Hausman, 2003) may be acute in the case of electronics, where industry changes are so rapid that substitution bias and particularly, new goods bias, may be particularly severe. Price indices are intrinsically designed to measure price changes in a bundle of goods, and poorly capture consumption changes or the dramatic market evolutions seen in electronic markets. More recently, Handbury (2012) uses income dependent preferences to find wide differences between low and high income consumers values from grocery markets in various cities. Her work has similar motivations to this paper, but like others in this literature, is very rarely informative of electronics, which has historically been a small portion of spending.

reasoning of how in practice the inequality could be either increased or compressed may be useful. Let us first consider a case where the benefits of innovation favor to lower income consumers. Consider two types of products: televisions and mobile phones. Primitive versions of these products were an extravagant expense at the time of their introduction. But as time passed, the cost of these products fell dramatically and today their ownership is fairly egalitarian across the population.[15] As a result of innovation, even low income consumers can purchase products that would be completely unaffordable and exotic only a decade ago. More generally, innovation may improve the purchasing power of important products for lower income consumers relative to higher income consumers.

On the other hand, firms are constantly producing improved and premium product models. Sometimes entirely new product categories are introduced, whose early models tend to be very expensive. Such innovation that often first emerges in higher end products would disadvantage low income consumers who tend to purchase products in lower price ranges. Additionally, a variety of reasons can cause older and less expensive products, and even whole product categories, to become less functional or obsolete. These reasons include the discontinuation of maintenance and compatibility, loss of network effects, or the "under provision" of durability, i.e. "planned obsolescence". The same previously mentioned products could serve as examples: late model televisions with high definition content are expensive, and require expensive programming, while mobile phones have evolved into versatile computers that can cost nearly four figures.

The topic of the paper is an unexplored empirical question. By many measures, a high degree of economic inequality exists in the United States today. However, the effect on inequality from these products has not been estimated. The question of this paper is to find evidence of this effect; the gains from these products could narrow inequality by being more egalitarian than other measures of inequality, or further increase it by overwhelmingly benefitting wealthier consumers. Even with a broad survey across the economic literature, I find little work that addresses this question.

[15] For the source, see Pew Research reports on cellphone and TV adoption

## 1.2 Model

As suggested previously, estimating consumer benefits from these products that are both durable and rapidly improving, requires an appropriate demand model. I use the structural dynamic model in Gowrisankaran and Rysman (2012), a model that allows for consumer heterogeneity. This model essentially nests a dynamic programming program inside the nested fixed point algorithm of Berry et al. (1995) (henceforth known as BLP). My innovation is measuring consumer welfare by income group. I setup the model first by showing the static utility of a given consumer, then show the dynamic framework and describe the population of consumers.

### 1.2.1 Per Period Utility and Consumer Heterogeneity

In each period (indexed by $t$) the consumer (indexed by $i$) enjoys a flow utility. Her flow utility depends on whether she decides to purchase this period. If she does not purchase, her flow utility is equal to the value of her current holdings of the good, or what she most recently purchased in a preceding period, plus an idiosyncratic type I extreme value error term $\varepsilon_{0t}$, that is *i.i.d.* across models, consumers, and time periods:

$$u_{i0t} = f_{i0t} - \varepsilon_{i0t} \tag{1.1}$$

(Note $f_{i0t} = 0$ when the consumer holds the outside good.) Otherwise if a consumer purchases, her flow utility is equal to the flow utility of the new purchased good (indexed by $j$ ), plus a price disutility term $P_{ijt}$ and another *i.i.d.* type I idiosyncratic error term:

$$u_{ijt} = f_{ijt} + P_{ijt} + \varepsilon_{ijt} \tag{1.2}$$

(In the case of linear utility in prices, $P_{ijt} = -\alpha_p^i p_{jt}$ so (1.2) is: $u_{ijt} = f_{ijt} - \alpha_p^i p_{jt} + \varepsilon_{ijt}$). The flow utility $f_{ijt}$ is a function of the characteristics of the product and the idiosyncratic characteristics of the consumer:

$$f_{ijt} = \alpha_0^i + \alpha_1 x_j + \xi_{jt} \tag{1.3}$$

Where $\alpha_0^i$ and $\alpha_1$ are demand parameters to be estimated, $x_{jt}$ is the observable characteristics vector of the good, $\xi_{jt}$ is the unobserved quality of the good. Note that in each time period each model has a new unobserved quality.

In the model, each consumer generally has different preferences and attributes. Importantly, each consumer has an income, $m_i$, and the distribution of incomes approximates the US household income distribution. Consumers are divided by their income into terciles, three equal sized groups corresponding to the lower, middle and high income groups in the United States. Let $L$, $M$, and $H$ denote these sets of consumers respectively, e.g. if $i$ is a low income consumer then $i \in L$.

Most parameters of flow utility are modelled as being constant across consumers. However, the constant term $\alpha_0^i$ and the price term $\alpha_2^i$ differ among groups of consumers or individual consumers. The constant term is modelled as a random coefficient, with independent normal distributions across consumers. Specifically, the coefficient of the constant term is:

$$\alpha_0^i \sim N\left(\bar{\alpha}_0, \sigma_0\right) \tag{1.4}$$

Where $\bar{\alpha}_0$ is the mean parameter and $\sigma_0$ is the standard deviation parameter to be estimated. The price coefficient does not have a random coefficient, however its mean depends on the consumer's income tercile:

$$\alpha_p^i = \begin{cases} N\left(\bar{\alpha}_p^L, \sigma_2\right) & \text{if } i \in L \\ N\left(\bar{\alpha}_p^M, \sigma_2\right) & \text{if } i \in M \\ N\left(\bar{\alpha}_p^H, \sigma_2\right) & \text{if } i \in H \end{cases} \tag{1.5}$$

The difference of the means is the channel through which I measure different income group's demand, and sensitivity to price. In general, these individual coefficients reflect potentially important differences in preferences that more realistically match the observed data. For example, lower consumer groups who are expected to have a higher price sensitivity will tend purchase and substitute to less expensive products.

### 1.2.2 Consumer Choices and Dynamic Structure

Consumers have an infinite horizon and use a common discount rate $\beta$. The data's frequency is monthly, and I set $\beta = 0.99$. In each period, the consumer observes available models, prices and both observable and unobservable characteristics (to the econometrician) of the product. Each consumer type has a holding of each model including the outside good, whose share is set to 1 at the first period of the model. In each period the consumer choices whether to purchase a product, and which product to purchase. If she purchases, her choice is one of the available models in the current period, $j \in J_t$. After purchases are determined, holdings of each model are updated at the end of each period.

The consumer's state variables are the vector of type I error terms for each product, denoted $\vec{\varepsilon}_{it}$, her endowment flow utility from previous purchases or the outside good $f_{i0}$, the characteristics of currently available models, and expectations of future models. Let $\Omega_t$ denote the current market state. $\Omega_t$ includes product information, such as number of models, attributes, prices and mean flow utility, but it also includes every other factor that influences future states of the market.

Therefore the consumer's state space at time $t$ is:

$$\{\vec{\varepsilon}_{it}, \Omega_t, f_{i0t}\}$$

Dropping the $i$ subscript and using the prime symbol to denote the next period value of an object, I can write the value function as:

$$V\left(f_0, \Omega\right) = \int \max \left\{ \begin{array}{l} f_0 + \beta E\left[V\left(f_0, \Omega'\right)|\Omega\right] + \varepsilon_0, \\ \max\limits_{j=1,\ldots,J} \left\{f_j - \alpha_p P_j + \beta E\left[V\left(f_{j,}\Omega'|\Omega\right)\right] + \varepsilon_j\right\} \end{array} \right\} g_{\vec{\varepsilon}}\left(\vec{\varepsilon}\right) d\vec{\varepsilon} \tag{1.6}$$

This equation is the consumer's infinite horizon expected utility. In any period, this is the maximum of two options in the current period: not purchasing and holding the current good, or purchasing a new good. Choosing to not purchase in this period results in the consumer receiving her endowment flow utility and the discounted expected value function in the next period. If she chooses to purchase, the consumer chooses the best option from her available choices, receives the flow utility of this good, pays the price term, and then receives the discounted expected continuation value with her new flow utility.

Unfortunately , equation 1.6 is intractable, as the state space of the products, and especially the space of information about future states is far too large. To make estimation tractable, I follow the previous literature by defining additional objects and making assumptions about consumer expectations.

Firstly, I define the *logit inclusive value* (IVS). This is a rewriting of the expected value of purchasing in a period, using the properties of the extreme value distribution:

$$\delta(\Omega) = \ln\left(\sum_{j=1,\ldots J} \exp\left(f_j + P_j + \beta E\left[V\left(f_j, \Omega'\right)|\Omega\right]\right)\right) \tag{1.7}$$

The logit inclusive value is the expected value of the consumer's best choice conditional on a purchase, and conditional on current holdings and $\Omega$. One can think of this as the value of purchasing, noting it includes the infinite horizon future values. Note that the use of logit aggregation in the definition of $\delta(\Omega)$ imposes a functional form on the value of new products, but this functional form is widely used in the related similar literature. Substituting this definition and applying the extreme value distribution to equation 1.6 yields:

$$V\left(f_0, \Omega\right) = \ln\left[\exp\left(f_0 + \beta E\left[V\left(f_0, \Omega'\right)|\Omega\right]\right) + \exp\left(\delta\left(\Omega\right)\right)\right] \tag{1.8}$$

Note that $\Omega$ only affects the value function through the current and future logit inclusive values, $\delta(\Omega)$. The consumer only uses $\Omega$ to make predictions of future values of the logit inclusive value. Next I make an important simplifying assumption about how consumers make predictions of $\delta(\Omega)$ that greatly improves the tractability of our problem. We assume consumers predict future values of $\delta(\Omega)$ using only the current value of $\delta(\Omega)$:

*Assumption* 1. Inclusive Value Sufficiency (IVS). Let $g_\delta$ denote the density of $\delta$. If $\delta(\Omega) = \delta(\tilde{\Omega})$ then $g_\delta\left(\delta\left(\Omega'\right)|\Omega\right) = g_\delta(\delta(\tilde{\Omega}')|\tilde{\Omega})$ for all $\Omega$, $\tilde{\Omega}$.

In other words, the logit inclusive value is a sufficient statistic for the distribution of future states of the logit inclusive value. This reduces the state space to two-dimensions, greatly improving tractability. However using the IVS assumption places restrictions on how consumers anticipate how current industry states evolve. For example $\delta(\Omega)$ can take on the same value because there are a large number of moderately valued products or a smaller number of high valued products. This poses a problem if there are multiple underlying industry states for similar valuations, and these underlying states result in different futures. However , this is unlikely in practice. Considering the enormous space in which the electronics industry can evolve, I observe that the industry tends to progress in an orderly manner: the number of available models increases, attributes improve and prices fall, and as a result valuations generally rise monotonically.

The IVS assumption allows us to describe a consumer's beliefs of future valuations of the market state as a scalar, and to approximate her beliefs of future valuations by her current valuation. Note that the assumption does not apply to consumers' ability to predict the population global valuation, nor is it a statement about technological change. The assumption is merely used to model the consumer's own beliefs of her future market valuations. The valuation, $\delta(\Omega)$, is determined by idiosyncratic consumer preferences over each product, and so both it and derived predictions are unique to each consumer. The assumption may seem restrictive, but I argue that it is reasonable to use a scalar to approximate perceptions of the future.

Indeed, it is plausible that consumers decisions literally involve using one dimensional, ordinal heuristics of future values. This compresses the myriad characteristics of the market into a workable approximation when making purchase decisions. A simple heuristic may be preferable over a more sophisticated predictions if future valuations of innovation are inherently unpredictable. If predictions are inaccurate, they are less valuable to produce, so consumers are more likely to use simple and less costly heuristics to anticipate the future. It seems plausible that information on future market states is limited because both the magnitude and set of innovations in these products can be so large.[16] [17]

In addition to the IVS assumption, I assume that consumers expect that logit inclusive values to evolve according to an AR(1) process. Consumers also have rational expectations, which means the parameters of the AR process are found

[16]Steve Jobs claimed: "You can't just ask customers what they want and then try to give that to them...customers can't anticipate what the technology can do. They won't ask for things that they think are impossible." (interview in "Inc Magazine", 1989).

[17]It is implausible for consumers to anticipate their valuation of new features (e.g., the introduction of voice commands on smartphones), or even the value from incremental improvements such as camera zoom or LCD size, since they have never experienced these innovations or combinations thereof, and also because the actual execution of the innovation is paramount (in fact the quality of execution can be unknown even to the product's designers). Additionally, little objective information is provided to consumers by producers by marketing. Instead, marketing strategies in these competitive industries gravitates to the two extremes of nebulous hype, or ostensible secrecy, neither of which seem conducive to producing accurate information.

using a simple autoregression:

$$\delta_{t+1} = \gamma_1 + \gamma_2 \delta_t + \nu_{t+1} \tag{1.9}$$

The specific form of an AR(1) process is not necessary. However given the reasoning above, this AR process is parsimonious and captures the important characteristics of the market, in that value of the market state is almost always improving. Note that in most situations, at the solution of the model the AR(1) process will generally predict an increase in the market state. However for extremely high valuations, AR(1) predictions will plateau at a steady state where predictions are on average not increasing. This is undesirable since we do not necessarily observe any plateau. Importantly, in practice the steady state is not reached; the economically significant values that consumers actually make decisions upon are below this plateau.

Using these assumptions we can now write the Bellman equation in am econometrically tractable form:

$$V\left(f_0, \delta\right) = \ln\left[\exp\left(f_0 + \beta E\left[V\left(f_0, \delta'\right)|\delta\right]\right) + \exp\left(\delta\right)\right] \tag{1.10}$$

### 1.2.3 Aggregation and the Criterion Function

The estimator is the standard GMM estimator used in BLP and subsequent related literature. It is defined as follows:

$$\begin{aligned} &\min_{\alpha,\sigma} \xi(\alpha,\sigma)' Z W Z' \xi\left(\alpha,\sigma\right) \\ &s.t: \hat{s}\left(\alpha,\sigma\right) = s, \hat{S}_m\left(\alpha,\sigma\right) = S_m \end{aligned} \tag{1.11}$$

Where $s$ and $\hat{s}$ is the vector of observed and predicted shares of each product; where $S$ and $\hat{S}$ are the vector of observed purchase shares of each income group ($\{S_L, S_M, S_H\}$) and their predicted counterparts; and where $Z$ is the vector of instruments (described below)and $W$ is the standard GMM consistent estimator of the optimal weighting matrix (see Hansen (1982)).

In order to estimate the model, I use a continuum of consumers with individual characteristics. As stated previously, each consumer has an income that is drawn from the US household income distribution, with a disutility from price whose mean is determined by the consumer's income tercile. In addition to this income heterogeneity, each consumer has another *i.i.d.* normal draw that modifies their constant term. As a result many objects in the model are unique for each consumer, and are appropriately indexed by $i$, i.e:

$$f_{ijt}, P_{ijt}, \varepsilon_{ijt}, \delta_{it}, V_{it}, \gamma_{1i}, \gamma_{2i}, \nu_{it}$$

Given values of these various objects, including the logit inclusive value, the Bellman equation, the flow utilities, and the AR parameters, I can write the probability of purchase of any product using the logit probability formula:

$$\hat{s}_{ij}(f_{i0t}, \delta_{it}) = \frac{\exp\left(f_{ijt} - \alpha_p^i P_{ijt} + \beta E\left[V_i\left(f_{ijt,}\delta_{i,t+1}|f_{ijt,}\delta_{it}\right)\right] + \varepsilon_j\right)}{\exp\left(V_i\left(f_{i0t}, \delta_{it}\right)\right)} \tag{1.12}$$

This is the probability of purchase of model $j$ in period $t$, conditional on the consumer's holdings and logit inclusive value. These purchase probabilities produce the expected market share in the given period as a function of the previous holdings of a consumer. Integrating over consumer types gives the total expected market share in a period given the vector of flow utilities $\vec{f}_i$ for each consumer.

This process produces predicted shares from a vector of flow utilities. Following BLP, to evaluate the criterion function (equation 1.11), I want to find the vector of mean flow utilities so that the predicted shares match the actual shares. To find this I use a (pseudo) contraction mapping similar employed by BLP. Firstly, define $F_{jt} = \bar{\alpha}_0^i + \alpha_1 x_{jt} + \xi_{jt}$, where I replace the individual random coefficient with the population average. Then I iterative calculate the following:

$$F_{jt}^{new} = F_{jt}^{old} + \phi\left(\ln\left(s_{jt}\right) - \ln\left(\hat{s}_{jt}\left(\vec{F}^{old}, \alpha, \sigma\right)\right)\right), \forall j, t \tag{1.13}$$

Where $\phi$ is a tuning parameter used in the computation. Note that other objects in the model, such as the EVs (1.10), logit inclusive values (1.7), and the AR consumer expectations (1.9), depend on the value of $F_{jt}^{new}$ (and on eachother). These objects must be calculated and solved for simultaneously during the computation of 1.13.

Finally, it is important to note that there is no proof that the fixed point of these functions is unique. By precedent, I use the same assumption in Gowrisankaran and Rysman, that given the average parameters $(\bar{\alpha}_p, \sigma)$, the model produces a unique vector of mean flow utilities $(\vec{F}_{jt})$. Note that by "average" parameter, I mean that I use the same assumption in the existing literature to assume there is a unique $(\vec{F}_{jt})$ given the *level* of the price coefficient over all consumers. However, given this level, I argue that I can identify the relative value of the coefficient among subgroups. Although I need to make additional assumptions for in the proof of identification, these are far more specific than assuming uniqueness. This strategy will be described below.

### 1.2.4 Instruments and Identification

Most of my identification strategy follows Gowrisankaran and Rysman (2012) and the preceding literature. As I do not model the supply side, I assume the characteristics of products are determined by a random exogenous process. Firms cannot commit to prices and set prices in each period in response to market conditions. Therefore, I assume product characteristics are exogenous but price is endogenous. To control for endogeneity, I use as price shifters the standard BLP style instruments that common in the literature. These are the product characteristics $x_j$, and statistics about competition in the same market (i.e. time period): the mean characteristics of competing models of the same brand, the mean characteristics of competing models of all brands, the count of models of the same brand, and the count of all other models of all brands. By being designed to capture the intensity of competition for every product, these instruments should therefore affect the prices and substitution across products.

I need to identify three types of parameters: mean coefficients on product characteristics, random coefficients (the standard deviation of the random coefficient), and the price coefficients that differ by income group. Mean parameters are identified by observing changing shares across market states. Random coefficients are identified in two related ways: by matching data to the models predictions of the dispersion in consumer purchases over a range of products, and secondly, by the model's dynamic structure.

Specifically, if there is a high degree of substitution between products with the same characteristic, the random coefficients on that characteristic will be larger than in the case where substitution is more uniform among products. This is the standard static identification of random coefficients. Additionally, the dynamic structure of the model further aids in identification, as random coefficients imply different consumers have different preferences and different holdings. For example, consumers with a high inherent demand for the product are represented with a higher constant term and will tend to purchase relatively earlier and have higher holdings in earlier market stages. As a result of these higher demand consumers being satiated to some degree, subsequent (and usually higher value) products will have comparatively lower shares. By taking this into account, identification is improved because flow utilities would otherwise be biased downwards in subsequent periods.

### 1.2.5 Identification of Heterogenous Utility Parameters

Letting demand be heterogenous by a consumer's income allows us to better evaluate the attribution of benefit to different groups. As described earlier in the model section, I assign different price coefficients (i.e. $\{\bar{\alpha}_p^L, \bar{\alpha}_p^M, \bar{\alpha}_p^H\}$) to each of the three subgroups of consumers in each income tercile. To identify these two additional coefficients, I use two moment conditions. These are the two ratios of the rate of purchases between the three groups, over all products, i.e. $S_M/S_L$ and $S_H/S_M$ (note that since the average is over all products each ratio is a scalar). Or equivalently, combining these ratios with the aggregate purchases of the three groups, these ratios imply an average share of purchases for each of the three groups. The model then matches these predicted shares with the observed[18] shares, i.e. $\hat{S}_m(\alpha, \sigma) - S_m$, where the preceding term is a vector of three elements. Since the shares of each income group are exactly identified, I do not use moments per se, but instead have the model match these rates exactly, analogous to how the BLP estimation matches predicted and observed shares.

As above, I rely on the assumption in Gowrisankaran and Rysman to identify the level of the price coefficient. However I show that, under certain conditions, given any one of the price coefficients, the two additional coefficients are identified. In other words, after I am given the level of the coefficient, I can identify the relative coefficients between groups. My proof employs strict monotonicity and two assumptions, one of which is fairly strong. Note that to simplify notation, I let $F$ be the vector of mean flow utilities for all products.

*Assumption* 1. Consider any solution to the model and let $\sum_j \hat{s}_{ij}(F)$ be the sum of the

[18]The rates of purchase by each income group are imputed using both the market and micro level data, as discussed in the data section below.

probability of purchases for any given consumer at vector of flow utilities $F$. Holding all parameters constant, for any $F'$ where $F' > F$ element wise, at the new solution $\sum_j \hat{s}_{ij}(F') > \sum_j \hat{s}_{ij}(F)$.

*Assumption* 2. Consider any solution to the model such that the following vector of inequalities by product hold element wise: $\hat{s}\,(F) < \hat{s}\,(F^*) = s$, where $F$ is the original vector of flow utilities and $F^*$ is the vector of flow at the solution, and where the strict inequality holds element wise. Then $F^* > F$ holds element wise.

Assumption 1 states that given the same parameters, strictly increasing the value of every single product will strictly increase the likelihood of the consumer's purchasing products on average. On the other hand, Assumption 3 says that strictly raising the predicted purchases of every individual product in the solution requires strictly raising the value of every single product. Although this seems intuitive, and can be shown if the products are substitutes, it may not hold in this model because products may be complements due to dynamic holdings. Specifically, increasing the value of one product in an early time period increases its holdings, which will depress purchases of a second generation of products in the future. But it is possible that the sales of a third generation class of products may rise due to the suppression of the second generation's holdings.

Although I speculate[19] that product are predominantly substitutes, assumption 3 is unlikely to hold in every circumstance. But, as will be seen in the proof, this assumption is exceeding strong for my purposes, as I only need flow utilities to rise enough on average that it increases average purchases for consumers. However, I believe making this assumption may make what drives the logic of the argument more transparent than assuming the weaker statement.

*Proposition* 1. Fix one of the three price sensitivity parameters $\{\bar{\alpha}_p^L, \bar{\alpha}_p^M, \bar{\alpha}_p^H\}$ Given assumptions 3 and 1, for each unique value of the vector $\hat{S}_m$, and other parameters, the remaining set of price coefficients is unique for any solution to the model.

*Proof.* Without loss of generality, let us fix $\bar{\alpha}_p^H$. Assume the model has reached some fixed point with parameters $\{\alpha, \sigma\}$. Consider a change in one of the two remaining mean price coefficients leaving the rest of the objects in the model unchanged (i.e. the elements in equation 1.2.3). To be specific, without loss of generality, let us perturb the equilibrium by having $\bar{\alpha}_p^L$ strictly rise to $\bar{\alpha}'^L_p$, (i.e. low consumers become

[19]It is likely that this second order suppression of demand is much weaker than the first one, and since every single product's share is being increased, this effect may be washed out in some sense.

more price sensitive). Now, remaining in this state, where the model is no longer converged, consider the sum of purchase probability of low income consumers (for brevity I henceforth will call this sum of probabilities “shares”). By the definition in equation 1.12:

$$\sum_{i \in L} \hat{s}'_{ij}(f_{i0t}, \delta_{it}) < \sum_{i \in L} \hat{s}_{ij}(f_{i0t}, \delta_{it}),\ \forall\ j \in J \tag{1.14}$$

Where $\hat{s}'_{ij}(f_{i0t}, \delta_{it})$ denotes equation 1.12 evaluated at the higher $\bar{\alpha'}^{L}_{p}$ but with all the other objects in the model unchanged. Also, since $\sum_{i \notin L} \hat{s}'_{ij}(f_{i0t}, \delta_{it}) = \sum_{i \notin L} \hat{s}_{ij}(f_{i0t}, \delta_{it})$:

$$\sum_{i \in L} \hat{s}'_{ij}(f_{i0t}, \delta_{it}) + \sum_{i \notin L} \hat{s}'_{ij}(f_{i0t}, \delta_{it}) < \sum_{i \in L,M,H} \hat{s}_{ij}(f_{i0t}, \delta_{it}) = \sum_{i \in L,M,H} s_{ij}\ \forall\ j \in J \tag{1.15}$$

Equation 1.15 states the predicted shares in this perturbation of the solution are lower than the observed shares. Now consider the new full solution to the model given the new parameter $\bar{\alpha'}^{L}_{p}$, denoting objects in this solution with an asterisk. Using the fact that the change in parameter only affects the model through the reduction of predicted shares, by assumption 3, every element of $F^*$ must be greater than $F$. Note the following equality must hold at any solution:

$$\sum_{i \in L} \hat{s}^*_{ij}(f^*_{i0t}, \delta^*_{it}) + \sum_{i \notin L} \hat{s}^*_{ij}(f^*_{i0t}, \delta$$

${}^*_{it}) = \sum_{i \in L,M,H} s_{ij}\ \forall\ j \in J$(1.16)Now assumption 1 implies: $\hat{s}^*_{ij}(f^*_{i0t}, \delta^*_{it}) > \hat{s}_{ij}(f_{i0t}, \delta_{it}),\ \forall\ j \in J$ and $\forall i \notin L$. Combining inequality with equation 1.2.5 yields:

$$\sum_{i \in L} \hat{s}^*_{ij}(f_{i0t}, \delta_{it}) < \sum_{i \in L} \hat{s}_{ij}(f_{i0t}, \delta_{it}),\ \forall\ j \in J \tag{1.17}$$

Therefore there is a strictly inverse monotonic relationship between price sensitivity and market share in solution for any change in one of the price parameters, so $S^*_m(\alpha, \sigma) \neq S_m$ does not hold.

Now consider the cases with a change in two parameters where the parameter change in opposite directions. Then and without loss, let $\bar{\alpha'}^{L}_{p}$ increase and $\bar{\alpha'}^{M}_{p}$ decrease. An argument identical as above shows that (1.17) still holds. Similarly, in the case where the parameters change in the same direction it can easily be shown that $S^*_H(\alpha') \neq S_H(\alpha)$. Having gone through all cases, we find that for any pair $\{\bar{\alpha'}^{L}_{p}, \bar{\alpha'}^{M}_{p}\} \neq \{\bar{\alpha}^{L}_{p}, \bar{\alpha}^{M}_{p}\}$, it must be the case that $S^*_m(\alpha, \sigma) \neq S_m$. Since we choose to fix $\bar{\alpha}^{H}_{p}$ arbitrarily, we have our result that given shares $S_m$, and fixing one price parameter, the remaining two parameters are unique.

□

The model is now formally identified. However, it is important to note that the model makes relevant internal predictions that are not clearly driven by empirical data. The moments currently only identify the different in average purchase rates by income groups, but say little about what the purchases are. Specifically, the timing of purchases for different income groups is not well identified and this may affect welfare calculations. This can be solved by adding additional moments that attempt to match holdings consumer behavior with these internal structure predictions.

## 1.3 Data Sources and Imputation of Purchases Rates

This paper primarily uses two different types of datasets. Firstly, I use three market level sales data sets for the United States for three major categories of electronics: camcorders, digital cameras, and DVD players. The market level datasets were acquired from NPD Techworld,[20] whose data is collected using point of sales data from electronic retailers or records of shipments from distributors, to create a comprehensive record of sales in the United States over this time period. Secondly, I use "ISMS Durable Goods Dataset 1", a micro-level panel dataset that contains observations at the sale level from a major big box retailer. This retailer dataset on sales was acquired from ISMS.[21] The panel dataset includes information on purchaser demographics, most importantly data on income, and also includes information on the purchases themselves, including price and category of product (but not the model of product). I discuss both these types of datasets I use in detail below.

For estimation, I will divide the dataset into three terciles, with the lower income groups defined as households with an income below $30K, middle income groups as households with incomes between $30K and $75K, and high households who have incomes above $75K.

[20]This consulting firm is a major source of retail sales data on electronics in this time. In general, NPD maintains relationships both with major retailers such as Bestbuy and CircuitCity, and major distributors collects point of sale data directly from company MIS departments. From this large subset of sales, NPD imputes the remaining industry sales to get an estimate of the entire market.

[21](INFORMS Society for Marketing Science).

I impute purchases of different income groups in the market level dataset by simply setting the relative purchases by different income groups to match the proportions in the retailer data. Specifically, assume there are three income groups: high, middle and low, and I observe each group purchases 50%, 30% and 20% of products respectively. Then for 10 million purchases, the purchases I impute to the groups are 5, 3 and 2 million respectively. This simple use of the micro-level data is simple and conservative, and I argue later, supported by the data. For example, this imputation would be valid if consumers of various income were on average equally likely to shop at big box stores for electronics. As a check, I use micro-data from the Consumer Expenditure Survey (CEX) to assess the validity of the retailer micro-data by comparing consumers from both datasets. This comparison is also below at the end of the data section.

### 1.3.1 Market Level Datasets

The market level datasets include all mainstream models for each of the three products over commercially important time periods. For camcorders, the data covers 2000 through 2006, which captures the market's infancy to its peak among mainstream consumers. The data on digital cameras is 1998 through the first half of 2001, a time period where digital cameras sales increased by a factor of 10, causing rapid decline in the film camera market. The DVD player data covers the period from 1997 to 2003, and like camcorders, spans the market's infancy to nearly its sales plateau.[22] Figure 1 plots the sales and price information from the dataset for all three product types.

The camcorder dataset contains 383 models and 11 brands, and monthly level data on prices and quantities sold of each model. The dataset also includes characteristics of the camcorders including size, mega-pixels, optical zoom and other features. The DVD dataset contains 523 models and products characteristics such as output information (composite, optical, Dolby Digital), MP3 and DVD-R compatibility, recording ability, and weight. The digital camera dataset includes 232 models and has characteristics such as pixels, zoom weight, and the presence of an LCD screen.

[22] Unit sales of DVD players and Camcorders level off around 2005 and these products experience a gradual decline as a fraction of consumer spending in the later half of the 2000s. Source: Consumer Electronics Association.

Since the goal of my paper is to assess the distribution of welfare over electronic products in general, it is important to note that the data on the three product types is plausibly representative of the electronics market as a whole. Each of these products have enjoyed widespread success in the United States and many other wealthy countries. As described, the datasets cover economically significant time periods with large growth in sales and rapid innovation that produces remarkable price declines. Furthermore, the nature of these products make them appropriate for estimating consumer demand in my model. These products are generally single function and tend to be used for personal purposes. Apart from perhaps DVD players, the products have relatively few commercial or academic applications. The single purpose nature of these products in this time period reduces the need to consider the proliferation of functionality in other roles and other complementarities. This is in contrast to products such as desktop computers, laptops or mobile phones, which can often be used for commercial purposes, and whose utility can be dramatically altered with new software and other products complementarities.

The data spans a time period that largely precede the rise of social media websites such as Facebook, YouTube, both of which became popular in the second half of the 2000s, and the data also precedes the proliferation of built-in wireless functionality such as WIFI or Bluetooth into electronics. As a result, there less concern about network effects, subscription or availability of services, such as wireless availability, when estimating demand for these products.

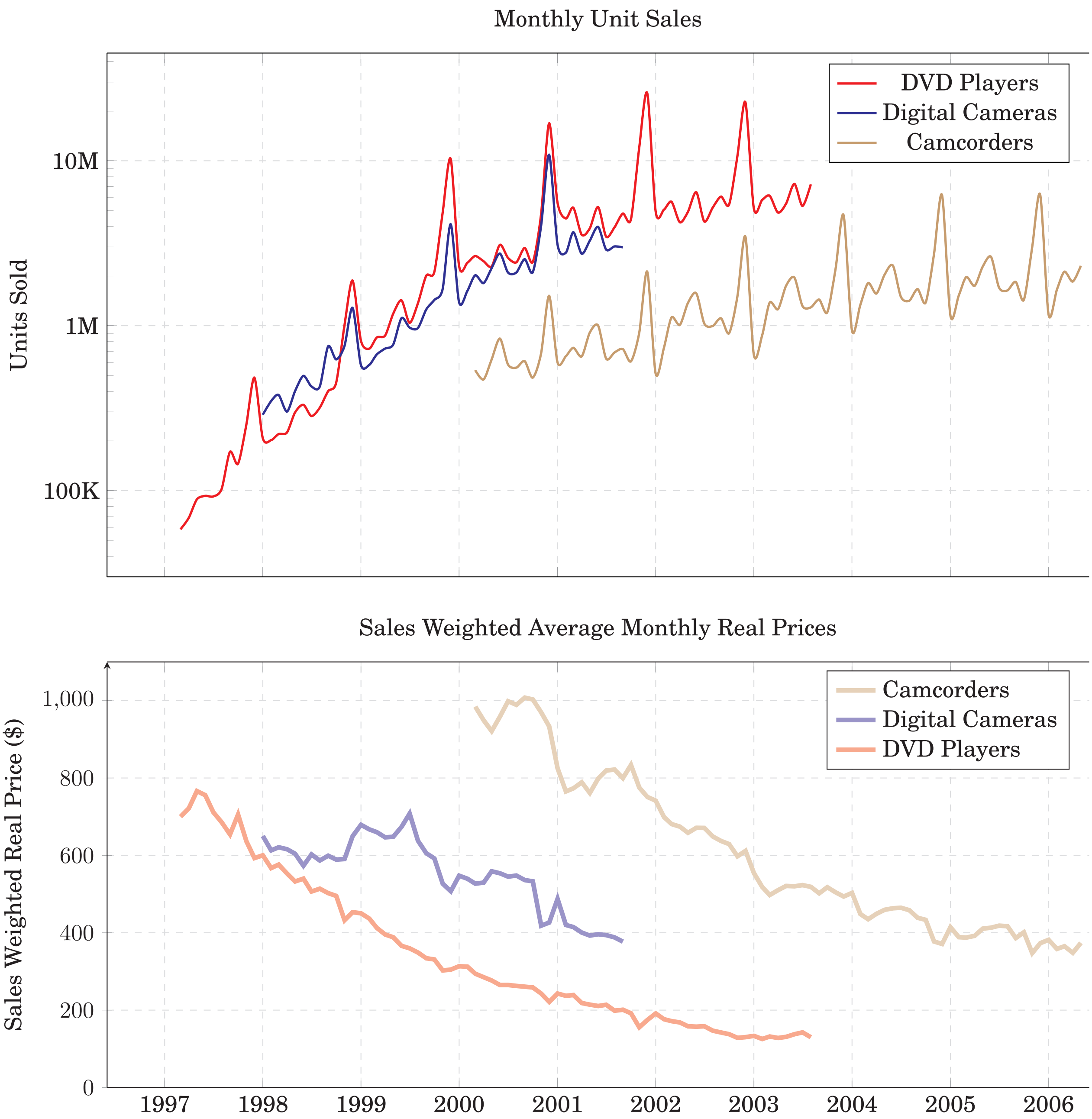
Before using the data to estimation, I drop observations where less than 100 models were sold in a month.[23] In the camcorder dataset, I drop products whose prices were less than \$100 or greater than \$2000, both to follow precedent and because these outliers may not be used for other purposes, for example, very expensive camcorders may be used by professionals.[24] I do not drop any outliers in price in the DVD player or digital camera datasets, as I do not have evidence of non-consumer purchases being predominant in any subset of these products.

Figure 1.2 shows sales and prices of all products in each of the three categories from the resulting data. The striking growth in sales is apparent as well as considerable price declines. Another important feature of the data are the massive sales spikes in

[23]With the exception for DVD players, where I drop sales below 1000. This further pruning of the DVD dataset is due to memory limitations in the computation for my hardware. Of the three datasets, DVD players have the largest number of sales, and the greatest number of models. Dropping observations of a model where sales were below 1000 units in a month reduces the dataset from 65 million units to 62.4 million units.

[24]The precedent is GR, who also performed the same price pruning.

Figure 1.2: Sales and Prices of Products




the holiday seasons. Estimating demand directly with seasonal spending is challenging, and the gains for estimation are unclear. I therefore deseason the data by dividing sales each month by a twelve constants that causes total sales in every month to be the same. Similar deseasoning is done in Gowrisankaran and Rysman (2012), who additionally estimate a structural model with 12 additional non-linear coefficients, and find very similar results to the deseasoned model.

As further checks for the importance of seasonality, in my microlevel data (described in the next subsection) I examine changes in the relative purchases for each income group by month. Relative purchase behavior by middle income groups remain the same. However, the relative share and total values of purchases by low income groups during the holiday spike fall moderately (by 3-7%), and the values for higher income groups increase by approximately the same amount. This reason for this trend is unclear, but one explanation may be that lower income groups spend relatively less on gifts of electronics during the holidays.

If so, I conjecture lower gift purchasing by lower income consumers could bias my estimates of inequality upwards by the following argument: If consumers do not enjoy surplus from such gifts, then including gift purchases will underestimate the shares of personal electronic spending for low income consumers, because a smaller proportion is on spending that do not provide surplus. However this depends on how we consider to be the value of gifts. For example, if gifts are given to persons of similar income, or purchased for other members of the households, the surplus of gifts should be included in the estimation.

Apart from this potential bias, which may be small, the seasonal trends do not seem significant. Finally, as further check, I examine prices and the number of models in each of the three datasets, and I do not find evidence for seasonal differences, which suggests that both the selection and nature of products sold during holidays are similar to other times of the year.

Table 1.1: Summary Statistics of Micro Level Data:
Overall Consumer Information

| | | Quantity of Purchases (Means) | | | | Spending |
|---|---|---|---|---|---|---|
| Income | N | All Products | Camcorders | Digital Cameras | DVD Players | Median Net ($) |
| Under $15k (8%) | 1,361 | 5.5 | 0.072 | 0.060 | 0.183 | 390 |
| $15k - $19k (3%) | 622 | 5.8 | 0.088 | 0.064 | 0.198 | 420 |
| $20k - $29k (8%) | 1,415 | 5.9 | 0.083 | 0.051 | 0.212 | 445 |
| $30k - $39k (9%) | 1,579 | 6.7 | 0.089 | 0.065 | 0.227 | 467 |
| $40k - $49k (9%) | 1,571 | 6.5 | 0.088 | 0.088 | 0.235 | 491 |
| $50k - $74k (22%) | 3,745 | 7.3 | 0.099 | 0.104 | 0.245 | 570 |
| $75k - $99k (15%) | 2,591 | 7.4 | 0.099 | 0.106 | 0.272 | 610 |
| $100k - $124k (8%) | 1,450 | 7.5 | 0.113 | 0.135 | 0.260 | 654 |
| $125k or More (14%) | 2,477 | 8.2 | 0.111 | 0.157 | 0.276 | 726 |
| Total (100%) | 16,811 | 7.0 | 0.096 | 0.100 | 0.243 | 546 |

In the description of the income brackets the values ending in "4k" or "9k" represent $4,999 or $9,999 respectively, i.e. "$19k" is $19,999.

### 1.3.2 Micro Level Data

The market-level datasets described above allow for identification of homogenous demand characteristics. In order to identify characteristics of demand that differ by the income of consumers, I use "ISMS Durables Dataset 1" from ISMS,[25] a micro-level point of sale dataset of purchases from a "big box" retailer[26] from 1998 to 2004. The panel contains observations of 173,262 purchases made by 19,936 households. This retailer dataset includes demographics such as gender, age, and geographic location of residence, along with point of sale, scanner-like information on purchases recorded at time of sale.

Importantly, the dataset also records each purchaser's income in one of nine groups. By using this grouping along with a key gained in communication with the NPD

[25](INFORMS Society for Marketing Science).

[26]There are over 1000 different retail stores observed in the dataset.

Group, the original source of the data, I can assign each household's income to fairly small income bracket. Most households are assigned an income, but some observations are missing income information. After dropping observations with missing income data, 16,811 households remain. Summary statistics of this resulting dataset are given in Table 1.1, which shows various measures of purchasing information of consumers, stratified by income group. There are similar tables in the appendix where the sample is restricted to households that purchase from each of the three product categories. These tables show similar results to Table 1.1.

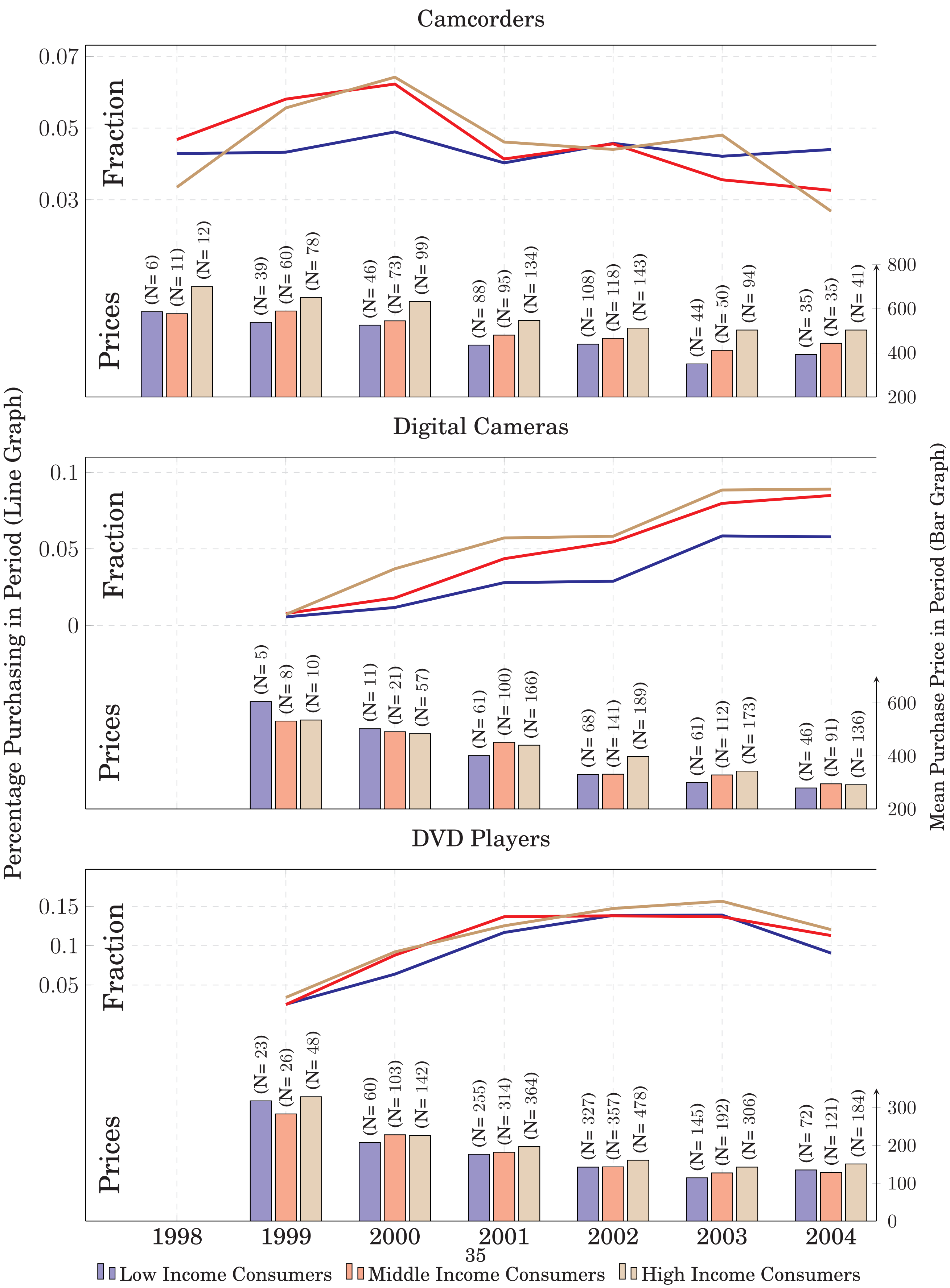
Although I ultimately group consumers to three income groups, all nine are displayed for illustrative purposes and to transparently assess the data's validity. For example, we would be concerned about bunching or implausible correlations with income that might suggest income groups may not have been recorded correctly. But no such problems are apparent in the data. As we might expect, a consumer's probability of purchasing a product of interest, her median spending and her number of purchases are all monotonic in income.

The data also indicates that consumers in the store level dataset who are purchasing products are a select subpopulation, whose spending on electronics is several times that of the average in the dataset, which in turn is higher than the spending for general United States consumers. Also, the distribution of incomes of all consumers in the retailer dataset is slightly higher than the US population as a whole (i.e. if we compared the histogram of incomes in the retailer dataset to census data, the retailer histogram would be to the right of census data).

Note that differences in characteristics of consumers between the retailer dataset and the US population as a whole does not invalidate the use of the retailer dataset as a proxy for income differences. I only need the condition that households of different incomes are equally likely to shop at the retailer, conditional on the decision to purchase. For example, if lower income groups are less likely to purchase electronics in general, we would expect to see a rightward shift in the income of consumers in the dataset, but the retailer data would still be a valid proxy for any purchasing behavior by income group.

Figure 1.3 plots purchase rates and average prices in the retailer data for each product, stratified by income group. This provides a useful comparison to Figure 1.2. To calculate the shares, the sum of product purchases are divided by the number of

Figure 1.3: Fraction of Consumers* Purchasing (Line Plot) and Mean Prices (Bar Graph with $N$), Stratified by Income



unique households observed in that period. Anomalously, purchase rates are not monotonically increasing. Particularly, the rate of consumers purchasing camcorders appears to fall over time. There is also a dip in the purchase rate of DVD players at 2004. These declines further reflect that consumers in the retailer data are a selected subpopulation. Perhaps consumers purchasing these products are switching to other venues, or alternatively the dataset at this time is adding consumers who generally purchase other products at the retailer. I conjecture the later is the case because of the secular changes in the electronics market as a whole, which widens the number of product categories available to consumers. To the extent that this biases the estimation, this is a concern if income groups differentially change their propensity to shop at the retailer. However, as I cannot evaluate this concern, I will assume that all income groups are changing their propensity to shop equally. Fortunately, when considering mean prices, the retailer data tracks the prices in the market level dataset. This supports the assumption that purchase behavior reflects purchasing by the entire population. In the next section I continue to investigate the validity of the retailer data.

### 1.3.3 Using CEX Data to Check Retailer Data Validity

By the 2010's, big box stores are an ailing business being supplanted by online retailers such as Amazon. However in the time period I examine is 10 years earlier, a vast length of time in this industry. In this period, big box stores still enjoy the lions share of electronic purchases, dwarfing online retailers such as Amazon.[27]

Nevertheless, an important question for my estimation strategy is to what degree the consumers in the micro level dataset of a big box store are similar to the purchasers in the US population as a whole. To address this, I use public microlevel data from the Consumer Expenditure Survey (CEX), which includes demographic information and self-reported purchases of various household products, including electronics, from a representative sample of households in the United States (after weighting). Using the CEX data, I compare the characteristics of consumers who purchase electronics to consumers of the same products in the retail data. 2002 is the earliest publicly available CEX data, so I examine purchases in the common time period of 2002 to 2004.

[27] Even by 2004, Best Buy's Revenue was still $25 billion, several times larger than Amazon's $7 billion. As electronics is only segments of Amazon's revenue, and a segment Amazon entered to later in its development electronics revenue only accounts for a fraction of this figure. Source: Public Investor Reports

Since the CEX dataset is designed to cover all major purchases of any household, not only electronics, categories in the CEX are coarser, and only a handful of these categories can be directly comparable to those in the retail dataset. However, I am able to match three product types: “computers”, “televisions” and “photographic equipment”[28] from both data sets in the same time period. To illustrate the characteristics of purchasers between datasets, I divide consumers into the three income groups as described above, and show the mean price of purchases and fraction of purchases made by each group. I also show calculate the mean of age and gender, which are demographic characteristics that are available in both datasets. Finally, for descriptive purposes, I show the total spending on the products implied by CEX weighting, stratified by income group. These results are shown in Table 1.2.

[28]“Computers” and “Televisions” both exist as categories in both the CEX and retailer datasets. However, while “Photographic Equipment” exists as a single category in the CEX data, the corresponding purchases retailer dataset are assigned to several finer subcategories: “Camera Accessories”, “Digital Cameras - SOHO”, “Film/Disposable Cameras”, and “Cameras”. To match the CEX data, I group these purchases together in the retailer dataset and restrict purchases to purchases greater than $50 dollars, which raises the resulting average price (to a more similar average price as the CEX data). The justification for restricting purchases to $50 and greater is to exclude small ancillary purchases such as memory cards and batteries, which are all recorded in the retailer data, but likely severely underreported in the self-reported CEX data. See Feenstra and Shapiro (2003) for a discussion of the differences and advantages of scanner data over survey data.

Table 1.2: Comparison of Consumer Behavior by Income Group
in Retailer Microdata vs. CEX Data

| | Income Group | Obs (N) | Mean Age | **Mean Price** | Female | **Share** | Pop. Spending |
|---|---|---|---|---|---|---|---|
| | | | | **Computers** | | | |
| **Retailer** | Low | 357 | 42.3 | **$ 732** | 0.42 | **0.313** | |
| | Medium | 357 | 48.6 | **$ 743** | 0.34 | **0.313** | |
| | High | 425 | 47.4 | **$ 794** | 0.33 | **0.373** | |
| **CEX** | Low | 1,748 | 45.1 | **$ 715** | 0.46 | **0.326** | $16B |
| | Medium | 1,809 | 43.1 | **$ 750** | 0.39 | **0.342** | $17B |
| | High | 1,881 | 45.5 | **$ 836** | 0.39 | **0.332** | $18B |
| | | | | **Televisions** | | | |
| | Income Group | Obs (N) | Mean Age | **Mean Price** | Female | **Share** | Pop. Spending |
| **Retailer** | Low | 1,215 | 45.3 | **$ 425** | 0.42 | **0.249** | |
| | Medium | 1,605 | 46.5 | **$ 510** | 0.35 | **0.330** | |
| | High | 2,050 | 48.7 | **$ 555** | 0.29 | **0.421** | |
| **CEX** | Low | 1,969 | 48.3 | **$ 268** | 0.53 | **0.440** | $5.6B |
| | Medium | 1,552 | 43.1 | **$ 293** | 0.44 | **0.316** | $4.5B |
| | High | 1,285 | 45.5 | **$ 426** | 0.42 | **0.243** | $5.0B |
| | | | | **Camera Equipment** | | | |
| | Income Group | Obs (N) | Mean Age | **Mean Price** | Female | **Share** | Pop. Spending |
| **Retailer*** | Low | 293 | 40.5 | **$ 258** | 0.39 | **0.194** | |
| | Medium | 497 | 45.0 | **$ 281** | 0.30 | **0.329** | |
| | High | 719 | 46.1 | **$ 313** | 0.30 | **0.476** | |
| **CEX** | Low | 665 | 46.8 | **$ 264** | 0.46 | **0.262** | $2.2B |
| | Medium | 814 | 43.1 | **$ 293** | 0.43 | **0.334** | $3.2B |
| | High | 1,011 | 44.8 | **$ 358** | 0.41 | **0.404** | $4.6B |

To generate the data in “Camera Equipment’ for the retailer micro-level data’, I group purchases from the following subcategories: “Camera Accessories”, “Digital Cameras - SOHO”, “Film/Disposable Cameras”, and “Cameras”, and restrict purchases to those with value greater than $50 dollars, The motivation is the suppress minor purchases in the retailer data that would be underreported in the CEX. The later is interview data that depends on monthly self-reporting, while the former uses records purchases with scanner-like precision.

With some qualifications, I argue the comparison supports the use of the retailer level data as a proxy of general purchase patterns. In the topmost section of Table 1.2, the characteristics of consumers purchasing computers in both datasets appear very similar. The purchase shares by income group, as well as the means of age, gender and prices match closely, espiecally considering the likelihood of typical measurement issues in both datasets.[29] Similarly, in the last section, the characteristics of consumers purchasing camera equipment appear comparable between both sources of data. This suggests consumers in the retailer dataset of these two important types of product appear to be representative of the population.

However, consumers purchasing televisions have very different characteristics. In this category, purchase prices in the CEX data are considerably lower than the retailer data. This is especially true for the purchases by lower income groups, who in the CEX data pay on average $268 dollars for a television, whereas their counterparts pay $425 dollars in the retailer data. Additionally, we see that lower income groups purchase televisions more than in the retailer data.

I argue these differences reflect the fact that the CEX data for televisions includes transactions in the secondary market that the retailer data omits. This explanation matches the discrepancies, since products in the secondary market will be less expensive, and it seems likely that lower income consumers more frequently purchase used products than consumers with higher income. But why is there a large secondary market for televisions? Note that the size of the secondary market is almost certainly proportional to the value of the stock of existing products. In the case of televisions, this stock may be large relative to the rate of new sales for idiosyncratic reasons. While televisions have experienced improvements and large price declines in recent years, the widespread adoption of this product greatly preceded other electronics (annual television sales exceeded 5 million by the early 1950s[30]) and unlike other electronics in the panel data, the rate of purchase has been fairly constant.[31] Televisions also seem likely to be less susceptible to damage or loss, compared to portable electronics such as cellphones or cameras. I argue these factors contribute to a larger and more valuable pre-existing stock of used televisions relative to new purchases, when compared to many other electronics products. This stock drives a potentially important secondary market for televisions that the retail data omits.

In comparison, products such as computers and cameras, that experience rapid improvements and dramatic increases in sales may have a minor secondary because rapid growth results in a smaller existing stock relative to current sales.

[29] Evidence of measurement error in the CEX has documented in Attanasio et al. (2004); Battistin (2003)

[30] For example, according to Retma's TV Factbook No.18, 7.3 million TVs were produced in the U.S. in 1953.

[31] The growth in television sales is unimpressive compared to many other electronic products. U.S televisions sales increased only from 27.1M to 35.5M units from 1990 to 1999 (Source: Consumer Electronics U.S. Sales). U.S. TV sales continued to grow in the 2000's to up to 39.1M in 2011. This was a roughly 45% increase over a period where population grew by 25%. Sales are now trending down with sales forecast to be 33M units in 2015. (Source: IHS iSuppli Research, October 2012).

Additionally, the large increase in sales itself suggests that the intrinsic values of older models are falling faster (or the value of new models rise faster), which further reduces the relative value of the stock through a second dimension. As a result, the secondary market may be small, and consumers in the retailer data could well reflect consumers as a whole. Nevertheless, the existence of the secondary market is important. Although I cannot measure such purchases, to the extent that a secondary market exists for the products in the market level data, my estimates may undervalue the benefits to lower income consumers. However, it is likely this underestimate is small relative to the total value of these products, for the reasons detailed above.

If there is comparatively little activity a secondary market for computers and cameras, we expect the purchasers to be similar in both the retailer and CEX dataset. Indeed, in Table 1.2, consumers purchasing computers and cameras have comparable characteristics. These rapidly improving products, computers and cameras, resemble the products in the market level datasets (camcorders, DVD players and digital cameras[32]). All of these product markets experience dramatic sales growth and will be much less affected by a secondary market. It is plausible that consumers in the micro-level retail dataset are representative of the population for my market level dataset, and I proceed with the estimation this assumption.

The proportion of consumers purchasing products is as follows in Table 1.3:

Table 1.3: Ratio of Total Purchases Relative to Low Income Consumers in Retailer Dataset

| | Income Group | | |
|---|---|---|---|
| Product | Low | Medium | High |
| Camcorders | 1.0 | 1.15 | 1.34 |
| Digital Cameras | 1.0 | 1.78 | 2.89 |
| DVD Players | 1.0 | 1.25 | 1.71 |

### 1.3.4 Descriptive Cost of Living Calculation Without Estimating Utility

Before presenting the main results from the structural model, generating a comparatively simple cost of living index may be useful illustrate the level of inequality we may expect to see, and to capture features of how the surplus changes

[32]The majority of cameras at the end of this time period are digital. Source: PMA Marketing Research, undated report, available at http://www.krlretirees.com/News_Items/20120901_US_Film_and_Camera_Sales_1995-2012_-_PMA.pdf. accessed 10/15/2014.

over time. Therefore, to give a preliminary description of the distribution of the surplus, I construct a relative "cost of living" index, without demand estimation, that uses the market level sales data, purchases of different income groups, and the canonical Laspeyres index:

$$\frac{I_{t+1}}{I_t} = \frac{\sum_j s_{jt} p_{j,t+1}}{\sum_j s_{jt} p_{jt}} \tag{1.18}$$

Note that while percentage changes in prices are obtainable from this standard chained index, creating a cost of living index is problematic because innovation (measured by price declines) is so large, and purchase rates rise dramatically. As a result, the standard Paasche or Laspeyres biases are exaggerated, but with opposite signs than normal since the price predominantly decreases. For example, consider a cost of living index that measures the change in income required to purchase some representative bundle of goods over time. If this bundle is chosen to match purchases early on in the market, where purchases are far fewer, the resulting measure of income will be an underestimate. Conversely, using a bundle of goods representative of purchases in later periods would produce overestimate of the surplus gains, because the number of purchases will be unrealistically large. This bias in magnitude is in addition to the problem that the bundle of products may differ significantly as time passes, due to large changes in product characteristics and prices.[33]

For these reasons, I calculate a unitless but cardinal "cost of living" measure derived from spending, and the price index that attempts to attribute the gains from the falling price index to each income group of consumers. Specifically, I calculate the price index by the standard chain weighted Laspeyres formula in each month. I then calculate the percentage change (usually a decrease) in the index, and multiply the negative of this percentage change by the total spending in the period. Finally, this series of changes is summed to the current time period $T$ to measure the total gains in surplus:

$$\text{COLI}_T = \sum_{t=1}^{T} \left[ -\left(1 - \frac{I_{t+1}}{I_t}\right) (s_{jt} p_{jt}) \right] \tag{1.19}$$

This provides a measure of the relative "cost of living" changes, or a measure of the magnitude of the benefits from price declines over time. This measure is homogenous among consumers of different incomes. As described above, this index is

[33]These problems stem from the fact that it is difficult to choose what is the appropriate bundle consumers would have chosen under different prices and different choice of products. This essentially is the motivation of this paper to produce an expenditure-like function to measure surplus.

proportional in both the amount of spending and the rate of price declines. To impute the benefits to each consumer group, I simply match the benefits to the proportion of spending. Specifically, I multiply this cardinal measure by the relative purchase shares of consumers, and normalize the resulting series by the level of the highest income group.[34] For example, imagine that the total measured gains were $10 and the relative number of purchases of high, middle and low income groups were 50%, 30%, and 20%, respectively. Then the gains attributed to each income group would be $5, $3 and $2. I plot this in Figure 1.4.

### 1.3.5 Calculating Consumer Surplus

As in Gowrisankaran and Rysman (2012), I calculate the value of these markets to consumers by measuring the estimated value of their flow utility or consumption from these problems. I find the dollar value of the subsidy or tax that makes their flow utility constant over time. Specifically, I create a subsidy term $S_{it}$ for each consumer in each time period such that :

$$0 = \sum_{j \in Jt} [f_j + P_j + L_{itj}] + S_{it} \tag{1.20}$$

Where $L_{it}$ is the logit utility value of choice, $-ln(1/S_{ijt})$. In the case of linear price utility, the utility value of a dollar is $\alpha_p^i$ so $S_{it} = s_{it} \times \alpha_p^i$, where $s_{it}$ is the value of subsidies in dollars and $\alpha_p^i$ is the consumer's price sensitivity which is composed of the mean price disutility in her income tercile and her individual price draw. I report the average $s_{it}$ over each individual in each income group such that (1.20) holds.

## 1.4 Results

### 1.4.1 Parameter Estimates

Figure 1.4 shows the parameter estimates from the structural model. The first group of estimates at the top of the table are the price sensitivity parameters. As we

[34] Note that this requires the assumption that the effect of the change in price of each product is distributed uniformly among consumers of different income groups, and that these consumers consume similar products.To the extent that low income consumers purchase lower priced products relative to other consumers, this approximation produces an is upward biased for the benefits to lower income consumers (such a bias makes the distribution more equal).

expect, lower income groups have a higher price sensitivity. Additionally, we see that the price sensitivity parameters estimates for digital cameras and DVD players are higher than that of camcorders. This is expected by theory: the sales of the former products are much higher than camcorders, in some cases by an order of magnitude. If wealthier consumers have a higher demand and lower price sensitivity, then it is case that sales of these products with a higher adoption rate must have travelled down the demand curve to consumers who are more sensitive to price.

The next group of estimates is the standard deviations of the random coefficients. The estimates of the standard deviation of the random constant coefficient are positive and in most cases, large relative to the constant coefficient, suggesting large variation in the valuation of these products.

Finally, the bottom panel contains the estimates of the coefficients on the product characteristics. Note that the values of these coefficients have no effect on the consumer welfare calculations for this paper. The welfare estimates only use the mean valuation of the product, net of price, while these coefficient estimates result from the model attributing components of this flow to various characteristics.

Digital cameras estimates are in the first row of the table. Here, most estimates follow intuition. Consumers value increased megapixels in the sensor, having more zoom range, and an LCD. In the case of weight and volume, the value of these characteristics are ambiguous in this market, so the model may have difficulties identifying this parameter. For many electronics products, consumers generally value smaller and lighter products, yet in the case of cameras, ergonomics and stability issues can make weight and size desirable; higher end cameras tend to be larger and heavier. Also omitted various bias may play a role, as many desirable characteristics of cameras are missing. For example, an omitted characteristic with great importance is optical quality. Another would be battery capacity measured by photographs taken. Given a fixed “technology level”, these omitted variables would be correlated with weight and size, due to physical engineering constraints.[35]

The second set of parameter estimates is for DVD players. The results here are mixed. A majority of coefficients are significant and follow economic intuition, but some estimates are unexpectedly negative. Consumers value optical outputs, Dolby Digital sound, VHS, recording and multidisk features. However, they do not value four characteristics: S-Video output, CDR-RW d MP3 functionality, which is hard to

[35]See various handbooks on cameras and photography for more detail, e.g. see Laikin (2010)

explain. Note that in the case of DVD players, all of these characteristics are dummy variables. As GR also found, the coefficients on dummy variables are harder to identify. This may be due to limited variation, as I observe that at many time periods, virtually every product in the entire market shares the same value of the characteristic.

The last set of parameters is for camcorders. Here most of the estimates follow our intuition. The exception is for nightshot, photograph capability and log zoom. Here again, the two dummies may pose challenges for the model. Log zoom may be negative due to an omitted variable bias: there is an tradeoff correlation between zoom and optical quality. Finally, note that based on (unreported) preliminary estimations, and GR's alternate specifications, log price fits the models slightly better than the linear price specifications shown, but linear price is much more more sensible when calculating welfare estimates as it provides a uniform utility value of money independent of purchases.

### 1.4.2 Welfare Results

Figure 1.4 shows the simple but descriptive "cost of living" index. There are several takeaways from the plot. Firstly, when considering the overall gains, the benefits to consumers rise most rapidly for from digital cameras, and still quickly for DVD players. The rise is almost exponential in these two cases, reflecting both the explosive sales growth and large price declines of these products. In the case of the camcorder market, the gains are not convex, which is attributable to the less striking, but still substantial sales growth in this market. Secondly, when considering the relative gains to different income groups, it is notable that low and middle income groups benefit similarly from these markets, with middle income consumers only slightly edging out the lower group. On the other hand, the high income consumers benefit significantly; in the order of greatest to least benefit of digital cameras, DVD players and camcorders. In the case of digital cameras and camcorders, with high income groups benefit approximately twice as much, while benefitting only 60% more in the case of camcorders.

It is notable that these gains are lower than inequality on consumption inequality, where higher income groups benefit approximately twice as much than the next

Table 1.4: Parameters Estimates from Model

| Digital Cameras ($\mathcal{N} = 2555$) | | DVD Players ($\mathcal{N} = 5928$) | | Camcorders ($\mathcal{N} = 4436$) | |
|---|---|---|---|---|---|
| **Price Sensitivity Parameters** | | | | | |
| Low | −12.7* | | −13.9* | | −5.070* |
| | (2.52) | | (0.808) | | (1.423) |
| Medium | −12.5 | | −11.6 | | −4.725* |
| | (7.57) | | (5.90) | | (1.07) |
| High | −12.3* | | −10.2* | | −4.774* |
| | (5.71) | | (5.38) | | (0.763) |
| **Standard Deviation of Random Coefficient** | | | | | |
| S.D. Constant | 0.008* | | 0.196* | | 0.199* |
| | (0.001) | | (0.001) | | (0.001) |
| **Other Parameters** | | | | | |
| Constant | 0.175* | Constant | −0.250* | Constant | −0.926* |
| | (0.026) | | (0.023) | | (0.122) |
| Megapixels | 0.161* | S-Video† | −0.095* | log Size | 0.021* |
| | (0.005) | | (0.017) | | (0.009) |
| Weight | −0.001* | Composite I/O† | 0.057 | log Zoom | −0.089* |
| | $(5.4e^{-4})$ | | (0.024) | | (0.009) |
| Volume | 0.401* | Optical† | 0.070* | log Pixels | 0.057 |
| | (0.032) | | (0.014) | | (0.098) |
| Zoom | 0.017* | Coax† | −0.016 | log LCD size | 0.089* |
| | (0.003) | | (0.016) | | (0.008) |
| LCD† | 0.227* | Dolby Digital† | 0.035 | DVD Media† | 0.362* |
| | (0.038) | | (0.019) | | (0.055) |
| | | DTS† | −0.002 | Tape Media† | 0.370* |
| | | | (0.019) | | (0.054) |
| | | CDR-RW† | −0.025* | HD Media† | 0.471* |
| | | | (0.004) | | (0.059) |
| | | MP3† | −0.181* | Lamp† | 0.067* |
| | | | (0.014) | | (0.008) |
| | | VHS† | 0.165* | Nightshot† | −0.084* |
| | | | (0.014) | | (0.008) |
| | | Programmable† | −0.036* | Photograph† | −0.063* |
| | | | (0.016) | | (0.022) |
| | | Recorder† | 0.406* | | |
| | | | (0.013) | | |
| | | Multidisk† | 0.168* | | |
| | | | (0.010) | | |

Standard errors in parenthesis. (†) indicates dummy product characteristics. (*) indicates Significance at the 5% level. All specifications include dummies of major brands, defined as top 10 brands sorted by number of models.

Figure 1.4: Relative Gains from Price Declines by Product

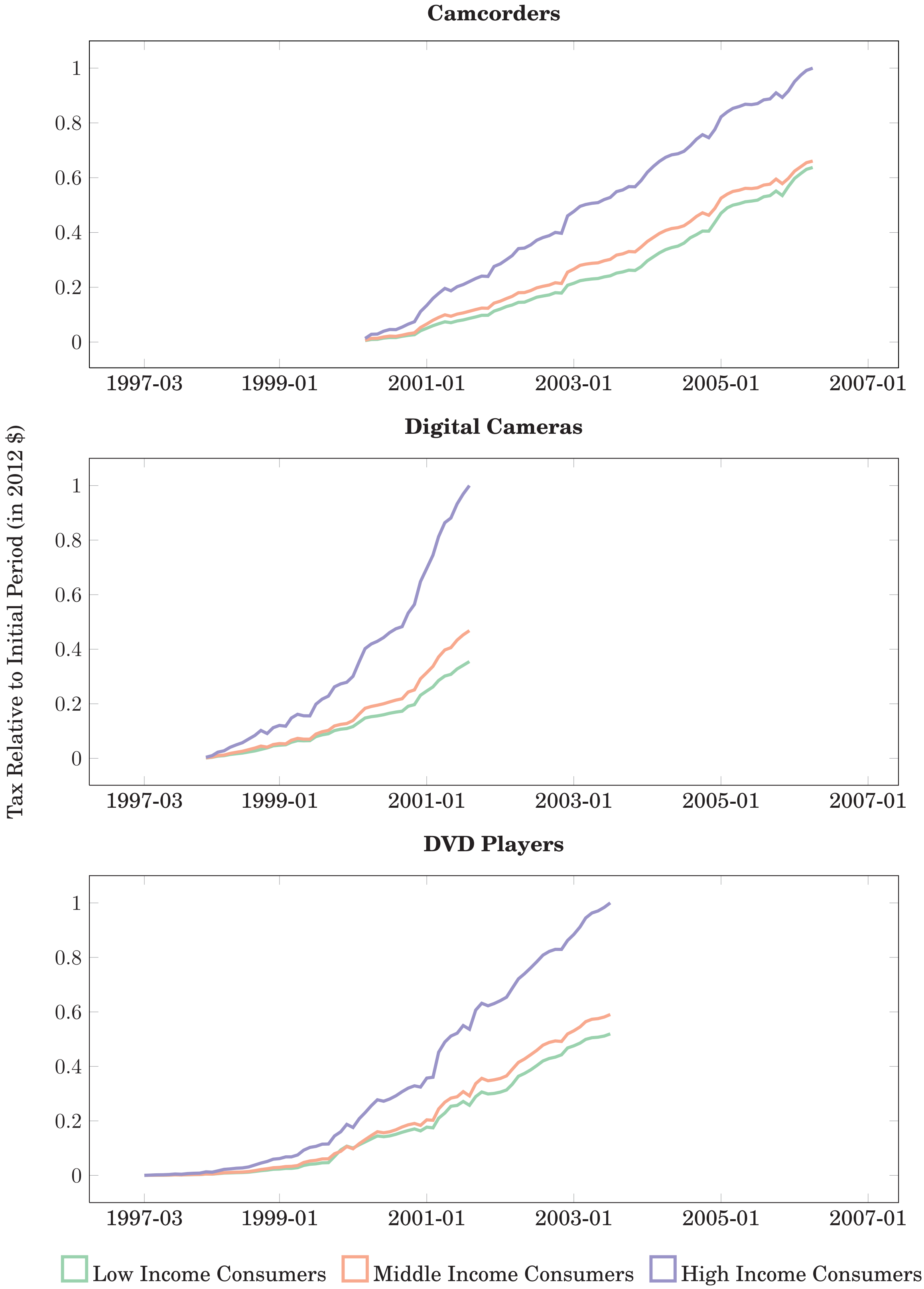
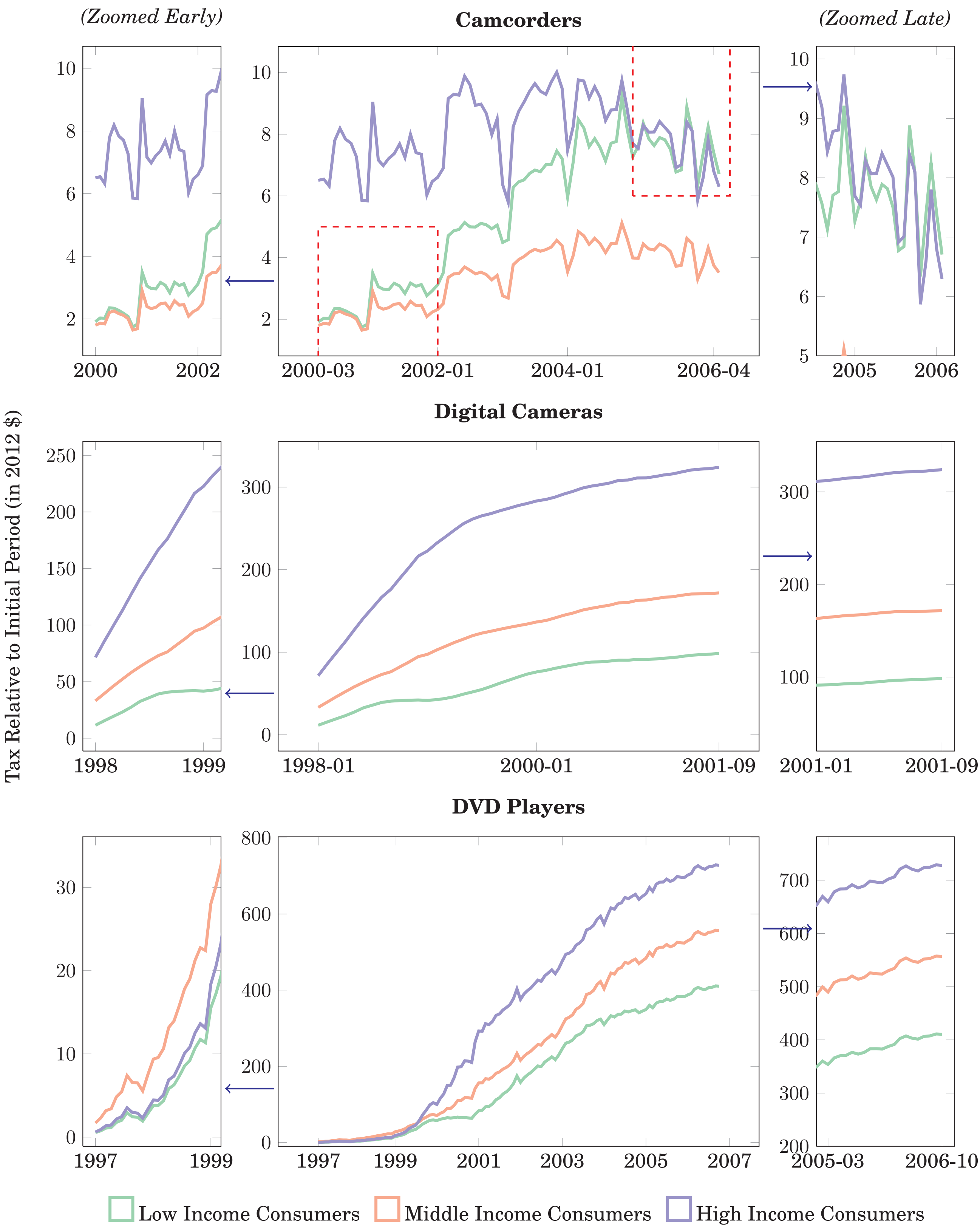


Figure 1.5: Welfare Changes by Product (Compensating Variation)



lower group,[36] which implies high income consumers have four times as much consumption; here I observe only twice the benefits. Also, interestingly, these comparatively egalitarian gains are most unequal in the markets that experience the most growth.

## 1.5 Discussion

The estimates of surplus from the three electronic products find gains in consumer welfare to be equal or larger size than the actual spending on these three products over the same time period. In the case of DVD players and camcorders, the surplus is equal to spending, while in the case of digital cameras the surplus is several times larger. This is likely related to the state of digital cameras at the end of the panel; compared to the other two products, the evidence suggests that the digital camera market is still relatively immature. The higher surplus may be due to consumer's beliefs about higher future states of the digital camera market. In the usual dynamic manner, consumers are withholding purchases of these products because of the high expected future states of the market. Current spending is suppressed while at the same time surplus rises for the same reason of that future market states will be much higher.

The high growth of the digital camera market may also be responsible its less egalitarian share of surplus. I conjecture that innovations are incorporated in high end product models, and therefore innovations tend to be available to high income consumers first. The digital camera market has not settled down, so the benefit to low income consumers is relatively low. Therefore I speculate this market will be more egalitarian in the future.

Generally, I find these estimates of the shares of surplus to be egalitarian. What qualities of electronics explain this? I argue that electronics tend to be egalitarian for several reasons. Firstly, note that like many other durables, few or often just one item of each of product is held at any one time. For the majority of consumers, there is rarely any reason to hold more than one digital camera, DVD player or laptop per person in any household. In fact operating multiple devices is often costly, both financially and in terms of the time and personal costs of learning the product. I.e.

[36] Note this is an approximation because the consumption inequality literature tends to use the ratio between the 90th to 50th and 50th to 10th percentile consumer etc., instead of comparing the gains from the terciles.

imagine trying to do work on multiple computers and managing information and compatibilities across devices. As a result consumers of higher means express their higher demand by purchasing more expensive products rather than more of them. Note that this is not true of many consumption goods. In many other industries consumers respond to price decreases by purchasing more items. For example, if all aspects of air travel became much cheaper, consumers would likely respond by purchasing more flights, instead of only purchasing upgrades. In a sense this reduces by one dimension the way higher income consumers can enjoy their higher income.

Secondly electronic products have a unique cost structure. I believe the main reason why electronics can become far cheaper over time is that the actual manufacturing cost of many electronics is small. Much of the costs of electronics come from research and development by companies, who extract profits by being market leaders. Therefore as time passes, competition and other forces (such as harvesting higher demand first) results in prices falling to very low levels. This is not true of most other products, including other durable goods. In a few years, some electronics have fallen in price by an order of magnitude, but automobiles will likely never fall in price by more than half in a lifetime. Physical materials and labor is large portion of the price of most conventional products, and these are difficult to reduce.

In 2014, consumers will spend approximately $213 billion annually on electronics. Performing a back of the envelope calculation, assuming that consumer surplus from electronics is the equal to total spending and using the average share of surplus, I find the average surplus for low income households to be $1,134. This is large, espiecally in comparison to the disposable income available to these households earning less than $30,000 a year (my definition of low income). A similar calculation finds that the surplus going to middle and high income groups to be $1,593 and $2,817 respectively. These figures are far more egaliatarian relative than what is implied by consumption measures of inequality, the most egalitarian measure of inequality.

## 1.6 Conclusion

Consumer electronics are a major source of consumer spending, with remarkably large and consistent growth in the United States. This paper develops and implements a strategy to examines consumer surplus for various income groups for three major categories of electronics. I investigate the distribution of consumer surplus from electronics, an important but a uninvestigated measure of inequality and I find that the surplus is fairly egalitarian: in each product, the distribution of welfare is more equal than other conventional measures of inequality, including consumption inequality.

With electronic products constantly improving and the types and categories of products continue to expand, there is every reason to believe electronics spending will only continue to grow, and along with it the importance of understanding its effect on the welfare of consumers. My results provide an illustration of the growth of welfare from electronics, and contribute to the understanding of inequality in today's society, where some measures of inequality appear to be at historic highs.

# Chapter 2

# Lasting First Impressions: The Influence of Initial Random Events on Long Term Decision Making

## 2.1 Introduction

> "I concluded that sociology was too hard, and returned, somewhat reluctantly, to economics" *–Gary Becker, "Lives of the Laureates"*

In 1950, a senior at Princeton University named Gary Becker was dissatisfied with his studies. Finding the topics in economics too narrow, he became interested in the field of sociology. To begin his studies there, he started reading a book by the influential sociologist Talcott Parsons. However, despite being highly recommended, there was a hitch in Becker's self-studies: Parsons' writings and jargon were impenetrable to Becker. Discouraged by sociology, though still ambivalent the discipline, he entered graduate school in Economics.

While Becker's initial foray into sociology was unsuccessful, his subsequent career in economics reflected his interests in the broader realms of social sciences. Perhaps had he been recommended another, more approachable text, Becker may have directly followed his interests in sociology rather than a PhD in economics. It seems possible that minor events, such as the recommendation of Parson's book, could have played a major role in Becker's life. In the same way, it seems possible that random idiosyncratic events may have influenced many of our preferences and decisions, possibly with lifelong effects.

Motivated by this concept, I use a novel dataset to explore how early random events may influence the decision making of individuals. I investigate how an individual solves the repeated problem of choosing among uncertain rewards, and I argue that minor exogenous shocks on early outcomes influence decisions made by individuals in the long term. When making subsequent decisions, individuals use information from the outcome of these early events, propagating these effects forward. For example, a one-time positive shock on an outcome from a choice causes the decision maker to repeat that choice. I find that these random shocks are used suboptimally as signals, in that individuals misattribute the effect of the random shock to the qualities of their choices. As a result, individuals may overweight early outcomes in their future decisions, making less efficient decisions ex post. The setting may allow this effect disentangled from competing explanations such as investment or experience. Instead, I argue this is based on the individual's perception of the outcome over other explanations.

Finding a non-experimental empirical setting to test these theories seems challenging. The random events should be unrelated to persistent characteristics such as ability, and also should not be separately observed by the individuals. In many settings, these conditions are hard to satisfy. For example, imagine the problems with randomly adjusting student grades to test the influence on their futures. Fortunately, with the growth of online activity and associated data, there are now newly available settings that provide data to examine my question. A unique setting that provides much of the structure is available in the form of an online multiplayer game, that both has the structure suited for this question, and which also provides a serendipitous wealth of data.

The setting is an online multiplayer game, called League of Legends, henceforth referred to as League. In this game, ten players in two teams face off in an arena, each interacting with the game by controlling their individual "champion." The decision I focus on is the player's choice of champion in these games. There are 124 champions, one of which each players must choose. Each champion is an in-game of avatar. Champions have very different characteristics and require different play styles. As a result, each player's performance varies widely between champions, and players tend to specialize in a handful of champions.

While there are measures of personal performance, in all matches the only outcome that affects one's place in the ranking is victory or defeat. This binary statistic is the only score that actually affects the ranked career of players, causing promotion or demotion. On average, with ten players in a game, one player can only be responsible for 10% of this outcome. Because players are randomly assigned to each team, much of a player's success is then the result of the random matchmaking outcomes.

My key finding is that the early outcomes with a champion, whether the game ended in victory or defeat, are associated with two later results in a career of a player. Firstly the likelihood of a player choosing that champion again increases with early victories. Secondly later performance and later victory is *negatively* associated with early victory, despite the fact that early personal performance shows a positive association with later performance and victory. I argue the best explanation for this is player misjudgement. I exclude other explanations with a variety of robustness checks to support the results.

## 2.2 Brief Background and Motivation

I investigate the degree to which relatively small circumstances can influence our first experiences with outcomes of choices. Random events and initial conditions influences many decisions, both large and small. For example, imagine the experience of one's first taste of coffee. The environment might influence the experience with the drink. A cup drunk on a leisurely afternoon with a friend would be different than having a rushed cup on an early workday morning. Also, most people have specific preferences over how they take their coffee. If a first time drinker found black unpalatable, many may not have the patience to try again. It is possible that for the same person, ex ante, these small events can put off, or attract them to an experience, and thus influences on their level of consumption for years. This paper investigates such random events on such decisions.

Perhaps in some ways early experiences are most influential for smaller decisions, since they are more numerous and the time and resources available to invest in decision making smaller. While I investigate a particular choice in an online game, I argue that the same mechanisms influence what books we read, the types of foods we enjoy, or the hobbies and friends we have. In a sense, perhaps these events have influence over the preferences of individuals. This paper follows behavioral economics.

This paper examines non-market based personal behavior using novel conventional data. This has precedent in several literatures that examine and test economic theory using a variety of real world settings. For example, expected utility and risk aversion aversion has been examined in game shows (Bombardini and Trebbi, 2012; Deck et al., 2008) and game theoretical concepts have been examined in professional sports (Walker and Wooders, 2001; Chiappori et al., 2002).

My setting notably differs than the topics above. Unlike the game shows, the motivations of players in my dataset are not financial, and while the stakes to individuals in a match are unclear, they are certainly lower than those in any professional athletic competition.[1] However, the smaller stakes seem more natural to my setting, and suited to an examination of learning and experimentation over a long period of time.

Furthermore, although it may seem strange at first, I argue that players in the competitive setting of the online game make decisions not dissimilar to important and common human tasks, such as working on a job. A career choice itself involves choosing between career paths, involving multiple jobs and so it is a repeated decision. These do not always involve high stakes: before committing to careers, people often make several small investments to see if they are a fit for the job, including shadowing, internship or volunteering. These activities, usually in a low pressure environment, involve assessing one's personal affinity for the work they perform.

In any given career, any workday can involves dozens of small decisions that are often repeated. Some, but few, of these decisions are for important stakes. When making these small decisions, most people are unlikely to be carefully ensuring every decision is optimal. Instead, they make many decisions based personal or idiosyncratic reasons. For example, these microdecisions could include the types of projects they choose to work on, which way to approach a new client, or what direction of literatures to research.

People also make many choices to satisfy their curiosity, and also act on selfless principles. In the same way that some gamers take their play very seriously, many successful people find enjoyment and a degree of playfulness in their serious work.

[1]There are professional League of Legends players with incomes over seven figures.

The microdecisions involved have individually low stakes, but could ultimately be of major importance.

## 2.3 The League of Legends Online Game

### 2.3.1 Overview

League of Legends is an online multiplayer game. It may be the most played such game of all time. The game involves players controlling characters called champions, who compete combatively in an arena. The game is developed and operated by Riot Games. In 2014, there were 67 million players active players a month, with over 27 million active each day.

Statistics about the player-base of League is limited, as Riot Games provides limited information about these statistics. However from an infographic in 2013, Google trends, and prize money, the interest in League is quite international. A few countries appear over-represented, but players are widely distributed across the world: in the Americas, many players come from Canada and the United States, Brazil, Chile, and Argentina. Sweden and Nordic countries appear to have a large player base in Europe and the Philippines and Australia in Asia and the Pacific. Other demographics of players are far less diverse: 90% of players are male, 85% are between 16-30, and the majority of players are attending or have completed some college.

In my dataset, I look at “ranked” matches, which is a category of matches that can only be played by players who have progressed after playing a fairly large number of matches, approximately 50-100. The other category of games are called “normal” games. There are over 12 million ranked players worldwide, which is approximately 10% of the total user base. Only ranked games are available through Riot’s API, and normal games are not in my dataset. Note that players are free to choose to play normal or ranked games, and generally alternate between the two modes. Because of this, it is possible that players put more effort and play more competitively in ranked games than they otherwise would, if they were restricted to a single game mode.

### 2.3.2 Champions

All gameplay during a match is through a player's selected champion. A champion is a digitally animated character, or avatar, in the game that exists in the game's arena. Champions have distinct identities and appear as humans, animals, or fantastical creatures, each with unique stylistic themes. With respect to gameplay, each champion also has unique attributes, including health, their offensive power, and manually activated abilities, that are equally distinctive. For example, there are fighter champions that rely on direct attacks, while others are spell casters with abilities that use on stealth and surprise to perform their roles. In fact, there are subcategories of fighters, and spell casters, and most champions have elements from multiple categories. As a result, champions have widely varying play-styles and require very different skills to master. Being good in one champion is not sufficient to be equally good with similar ones. With over 120 champions, few if any players are equally skilled in more than a handful.

### 2.3.3 Organization of Matches and Gameplay

To play a match, players enter a blind queue to join a pool of other players waiting for a match. In the pool, players have little control over the composition of their teams and never have any information about the players in the pool.[2] After sufficient players are in the pool, players are matched using an algorithm into two teams of five players that are approximately balanced by ability. The algorithm does not try to design teams to associate certain players, nor does it appear to take into consideration the identities or traits of the players. In other words, the algorithm appears to be one consistent with a system that uses a single scalar variable to match players by ability. There is evidence for this. For example, sometimes the system fails after forming teams for a prospective match. After quickly rematching, many of the same players are matched in the new prospective match, but the players in the original teams are scrambled across the two new teams.

After being matched into two teams of five players, players converse with teammates to jointly make several strategic choices that include deciding upon the roles each player has in-game. Teams self-organize into approximately five distinct roles, one

[2]Note that players can queue with a friend, and in this one way can control the identity of one player on their team. However, the algorithm takes this into account to match teams so that each has an equal chance of victory.

for each player, in particular areas of the arena. After choosing a role, players then choose a champion that tends to suit their roles. Although never explicitly prescribed, teams follow common conventions so that usually one of each of various canonical roles are represented on a team. This is similar to how there are codified roles in various sports, such as a center in hockey, quarterback in football, or setters in volleyball, that exist without explicit rules. These conventions are deemed optimal for any match, and are generally followed with only minor modification.

Because of the choices of other players, each player has considerable, but far from complete control over which champion they ultimately play. Random conflicts can occur. If by chance two players are matched to the same team who desire the same role, one generally defers to the other. If neither player defers, effectively a lottery will usually award one player the oversubscribed role, and the other player will accept another role. Because most champions are deemed optimal with a specific role, a player will then choose a different champion as a result of this conflict, which is the result of random matchmaking. Also, only one instance of each champion can be chosen in a match in ranked game. Therefore players are denied their choice if their teammates or opposing team chooses their champion. Finally, each team has the option of vetoing several champions, so that they cannot be selected. For all of these reasons, I estimate that less than half of the time a player can expect to have a particular champion they desire.

Once the match starts, activity in the game involves increasing the player's champion's power by gathering resources, fighting enemy players, and obtaining objectives. While statistics are available that reflect all of these activities, in this paper, I focus on the statistics associated with fighting: kills, deaths and assists. Kills are awarded when a player's champion damages an enemy champion so that their health falls to zero. Assists are generally awarded for damaging an enemy soon before their death, but not dealing the killing blow. Deaths, self-explanatory, are usually caused by enemy action. Kills and assists have almost always an absolute positive impact on the game, and deaths have an absolute negative impact. Although these events have important effects on the outcome of the match, they are not recognized by the ranked system, which only values victory or defeat. Despite the fact that these statistics are permanent, and often a focal point of criticism or praise of a players during and after a game, they are not believed to affect subsequent matches in any way.

## 2.4 Data

I examine player match data history from the official league of legends API. The dataset consists of match level observations of all ranked matches for 63,346 players from March 30, 2014, to May 15, 2015, from the League's North American servers. There are 29,078,519 matches in the dataset. Each observation of each match contains its outcome, victory or defeat, and also measures of the player's personal performance including "kills", "deaths" and "assists". As mentioned above, other game modes called "normal" are played, but these are not available. Ranked matches are a subset of games played for status in league of legends. The raw dataset is processed by dropping all observations with a champion a player has played less than 30 games. I also remove a subset of matches related to creating a control described in the robustness section. This processing reduces the dataset to a size of 12,037,660 observations.

There is no mechanism available in the API to randomly select players into the sample. To construct the dataset, a crawling strategy was used: starting with a seed of players, I find a list of matches by these seed players. Looking at the new players in these matches, I in turn find the players who played with them, etc. This will over-sample players who play more games. Fortunately, because matchmaking randomly matches players, this should otherwise provide a representative sample.

The focus of the paper is the influence of outcomes in the initial matches a player plays with a champion on the outcome and performance statistics in subsequent matches in a player's career. The average of the independent variables of interest: victory, kills, deaths, and assists, are calculated over the first eight matches a player plays with a champion. Outcomes for subsequent matches with the same champion are regressed this average. "Early" is used when referring these variables. For example, "early win rate" and "early kills" refer to the average victories and kills over the eight match period bin.

Table 2.2 gives descriptive information about the choices of champions of players in the total dataset. There are 63,346 players. On average, players have played large number of champions: over 40. Although players try a large number of champions, they have strong preferences of what champions they play and on average, players play their most popular champion over a fifth of the time. Their next two most popular choices take up around a tenth of their games.

Table 2.1: Summary Statistics on Processed Data

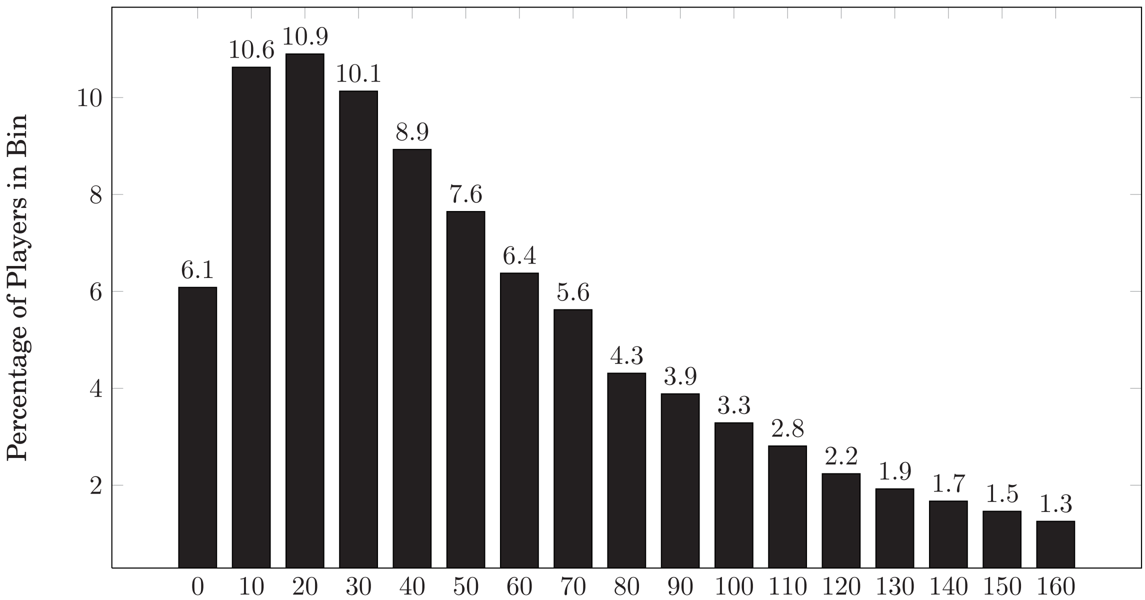
| | Mean | S.D | Min. | Max. |
|---|---|---|---|---|
| *Performance in First 8 Matches* ($N = 1{,}667{,}520$) | | | | |
| Winner | 0.52 | 0.50 | 0 | 1 |
| Kills | 6.44 | 4.91 | 0 | 49 |
| Deaths | 6.30 | 3.39 | 0 | 41 |
| Assists | 9.39 | 6.15 | 0 | 54 |
| *Performances in All Matches 11 and onward* ($N = 12{,}145{,}741$) | | | | |
| Victory | 0.53 | 0.50 | 0 | 1 |
| Kills | 6.91 | 5.12 | 0 | 58 |
| Deaths | 6.38 | 3.45 | 0 | 42 |
| Assists | 9.58 | 6.28 | 0 | 61 |
| *Other Statistics* | | | | |
| Total Matches Played | 111.09 | 132.21 | 30 | 2158 |
| N | 12,037,660 | | | |

Figure 2.1 is a histogram of the count of matches for each player's most played champion. The career of most players with any champion is under 50 matches, but there are a substantial number of players who have played 100 or more matches with a champion.

Table 2.2: Champion Choices by Players on All Data

| | Mean | S.D | Min. | Max. |
|---|---|---|---|---|
| *Fraction of Games By Most Played Champions* | | | | |
| First Played | 0.21 | 0.14 | 0.021 | 1.0 |
| Second Most Played | 0.12 | 0.06 | 0.001 | 0.5 |
| Third Most Played | 0.09 | 0.04 | 0.002 | 0.3 |
| *Other Statistics* | | | | |
| Different Champions Played | 43.38 | 26.61 | 1.000 | 124.0 |
| Total Matches Played | 459.04 | 397.52 | 1.000 | 4811.0 |
| #Matches/#Champions | 0.04 | 0.06 | 0.008 | 1.0 |
| N | 63,346 | | | |

Figure 2.1: Proportion of players who have played with at least one champion for the indicated number of matches (N = 63,346)




X-axis indicates lower bound of bin (i.e. “10” indicates bin with 10-19 matches played)

## 2.5 Theory and Estimation

### 2.5.1 Theory

The decisions players make over champions resemble a class of multi-armed bandit problems. In the standard multi-armed bandit problem, the goal of an agent is to optimize their payoffs over several periods. In each period, they choose between mutually exclusive "arms" or choices, each of which results in a stochastic reward, which are independent across arms. Arms have states, and choosing an arm changes the state and the future payoff of the arm. However, the state of an arm only changes when chosen and are not affected by the choices of other arms. Generally, the optimal strategy involves the balance of "exploration", where agents make choices to obtain information about those choices, and "exploitation", where agents choose their best expected option with the information they have. To be more concrete about the model, in Appendix A, I present more formally the standard multi-armed bandit problem.

The multi-armed bandit problem and the players decisions over their champions have several similarities. In both, an agent is selecting one of several choices (champion) to maximize an outcome (victory or enjoyment) in discrete time (matches). It seems intuitive and consistent with the data that players explore different champions, but ultimately choose their best champion.

However the two settings also differ considerably. For example, in the classical multi-armed bandit problem, the outcome of any choice is independent of the history of selecting other choices. In this setting, this implies that a player's outcome with a champion must be independent of his history of playing all other types of champions. This is quite unrealistic, because there are common skills involved in the play of all champions, and also players generally learn about other champions during play. Also, while players are motivated by victory, players might also wish to entertain themselves by choosing other champions, despite this choice being suboptimal. Additionally, unlike the classical bandit problem, players are also making an investment problem: their play with champions improve as they play more games.

Because even the simplest and most stylized multi-armed bandit problems have complex analytical solutions that are extremely difficult to compute, their systematic use by any player seems unlikely. For these reasons, and also because a formal model with the extensions above is beyond the scope of this paper, I do not present a formal model for the behavior of players. Instead, I point to the existing literature on bandit problems to heuristically argue several predictions.

First, when a choice results in good outcomes, players are more likely to make that same choice in the future. This is justified in theory by the analytical solution to bandit problems, as shown in Gittins and Jones (1974). Essentially, their solution is to compute an index for each choice, and the optimal solution is to choose the highest index at each time. This index increases in expected reward. In turn, expected reward increases with payouts using the process of Bayesian updating, it follows that players should be more likely to optimally choose champions with which they experienced good past outcomes. Evidence of individuals using Bayesian updating has been found in experiments such as Acuna and Schrater (2008).

Secondly I make the less intuitive prediction that early win rate is associated with *lower* subsequent player performance. Note that this is despite the fact that underlying player skill implies there should be positive correlation of outcomes across time: players who are more skilled are likely to more successful in both early and later games.

The possibility that players can make suboptimal choices has a theoretical justification in Rothschild (1974). His interesting result shows it is possible for rational agents to settle on a suboptimal choice. The intuition behind Rothschild's result is that in balancing exploitation and exploration of choices, the amount of exploration by agents is finite. Agents employ a cutoff rule whereupon they abandon a choice if they receive sufficient negative information from their past play. It is possible that by receiving enough negative random draws from a choice, they will reach this cutoff and never explore it again, and instead settle on the remaining options. Rothschild shows it is possible for agents to abandon the optimal option. To my knowledge, there are few empirical studies of bandit problems, and all are experiments, and none are related to Rothschild's result.

### 2.5.2 Estimation

My estimation strategy uses the argument that, because of underlying ability, early win rate is expected to be positively associated with later performance and victories.

If a negative association is found, I argue this is due to suboptimal player choices.

To examine the changing effect of these early variables of interest on later matches, I divide the subsequent matches into bins containing 20 matches, and the bins beginning at match 11. So the first bin contains statistics from matches 11 through 30 where the player plays that champion. Examining these partition of a player's career allows us to observe trends, and resolve effects that may be masked or net out, such opposing counterbalancing negative and positive results. I choose the bin of matches 50-70 to balance the longevity of results with statistical power.

I examine this association in my key specification of early win rate with later win rate. The empirical framework is as follows. There are $i \in I$ players who may play one $j \in J$ champions in a match. The observations are at the player-champion level, and each observation contains information about outcomes and performance of that player playing the champion. The estimating equation, estimated by linear regression, is:

$$Outcome_{ij} = \beta_w EWR_{ij} + \beta_c Controls_{ij} + \beta_f CWR_{ij} + Champdum_j + \varepsilon_i \tag{2.1}$$

where $EWR_{ij}$ is the early win rate in the 8 first games played by player $i$ with champion $j$. $Controls_{ij}$ include means of measures of player performance in the the same first 8 match period: the deaths, kills, and assists. The controls also include the average win rate of the player over all his or her champions. The specification also includes dummies for each champion, $Champdum_j$.

$CWR_{ij}$ is the control win rate variable used in some specifications to calculate the win rate of the individual when using *other* champions ( i.e. $J \setminus j$). This average is calculated in the same way as $EWR_{ij}$, also using the other 8 matches, starting at the time of the first match of champion $j$. The time periods of the control win rate variable closely overlaps the time period $EWR_{ij}$ is calculated.

The purpose of the control win rate variable is to isolate competing explanations for effects we find from the variable of interest, early win rate with a champion. If another explanation was responsible, for example, if players subsequent performance improved because of increased motivation after a series of defeats, or they were matched with weaker opponents, the resulting effect should be seen in both control win rate and the win rate of interest. If there was a difference in the results from the control win rate and win rate with a champion, it seems plausible to ascribe this to the champion the player uses.

Standard errors are clustered per player, and specifications generally include player fixed effects.

## 2.6 Results: The Impact of Early Events on Long Term Outcomes

In the results, I find strong evidence that early success with a champion leads to poorer later outcomes, consistent with my theory that players are "unduly" influenced by early random outcomes.

Table 2.3 contains the key results of the paper from regressions that examine the influence of early game outcomes. Most specifications have the dependent variable as the average win rate for the frame of matches 50 through 70, while the last specification in column 5 looks at matches 11-30. Column 1 shows a simple regression of the independent variable, the outcome of a match with that same champion, on the dependent variable, early win rate with a champion. Here the coefficient of the early win rate is positively correlated with the win rate with a champion. This is intuitive and consistent with the possibly that ability, an omitted variable, is correlated with both early and late success.

Indeed, when controls for champion performance are introduced in column 2, the effect of the dependent variable is reversed. We also see that the performance variables, kills, death and assists, show their natural associations with later victory: the number of kills and assists are positive predictors, while the number of deaths is a negative predictor. These are the expected effects if early player ability was a predictor of future ability. However, the variable of focus, the early win rate with a champion, now shows a negative relationship with later champion win rate. So, as is consistent with my hypothesis, once controls for performance are introduced, we see the hypothesized negative effect. Note that column 2 is the preferred specification.

Also, the coefficient of the effect of the early win rate on later win rate is -1.04%. If this is causal relationship, if we could change a single early loss to a victory, or made a 0.125 point increase in the win rate, we will observe a 0.13% decrease in the win rate, even fifty games later.

Column 3 and 4 of the same table show alternative specifications. Column 3 does not include fixed effects, while column 4 examines the win rate of other champions. Removing fixed effects removes significance on our variable of interest, although the sign remains negative. The coefficient on our variable of interest remains largely unchanged. If the effect of the early games was mediated through intermediate performance, we would expect adding these additional control win rates to change the effect.

Finally, column 5 runs the preferred specification on the 10th to the 30th match played with a champion instead of the matches 50-70. This bin of matches is much closer to the early period where the early outcomes and control variables occurred. Because there are more observations when we examine matches where the player has used the champion less, more statistical power is available. Here, the coefficients confirms the results in the preferred specification previously discussed in column 2. Looking at kills, deaths and assists, we now see statistical significance and the expected relationship of all of these performance variables on later outcomes. All coefficients also have larger magnitudes. We see that the early win rate now has a larger effect of -2.20, twice as large as it was for matches 51-70 in our preferred specification.

The regressions in Table 2.4 uses similar specifications as those in Table 2.3, but the dependent variables are the measures of player performance in later matches. Therefore unlike Table 2.3, the dependent variables changes by each column. Again, the matches I examine are those 51 through 70. Results in columns 1-3 look at the effect of the early win rate on the three performance variables deaths, kills, and assists, respectively. In these simple specifications without controls, the only significant association is a negative correlation of early win rate on deaths. In this one significant result, early success is correlated with a good outcome.

Columns 4-6 of Table 2.4 look at the same dependent variables, but with measures of early performance added. Just as in Table 2.3, the addition of the performance controls dramatically change the coefficient on the variable of interest, early win rate. We again see the effect of early success predicting poorer later outcomes. For example, in column 4, where later number of kills is the dependent variable, we see the coefficient early win rate is associated with fewer kills. In contrast, the controls of performance generally have the natural associations: more kills and assists has a positive influence on later kills.

In column 5 and 6, we see similar results to column 4. Early win rate is negatively associated with later outcomes: there are fewer kills and assists with that champion later in a player's career. In contrast to early win rate, when looking at controls, we generally see the natural associations. Early kills are positively associated with later kills and negatively associated with later deaths. The converse is true of early assists: they are negatively associated with kills and deaths and positively associated with assists. Early deaths are positively associated with both later deaths and kills. Some of these signs are counter intuitive such that kills are a negative predictor of assists, however this is explainable by variation in styles of play in which deaths and kills are highly correlated.[3]

In addition to the tables which focus on several coefficient for a few bins of matches, I also include a plot which shows the coefficient of interest from over various points in a player's career over the entire sample. This is shown in figure 2.2 . This illustrates the level and trends of the coefficients of interest on early win rate on various statistics over a player's career. Specifically, I take the preferred specifications (column 2 of Table 2.3 for early win rate and columns 4-6 of Table 2.4 for other performance variables) and run them the entire sample. The coefficient of interest, early win rate, is interacted with 10 dummies, one for each of the 20 match bins from 10 to 230. I also include these dummies. The coefficients of the early win rate on these dependent variables are then plotted for the means of our four variables of interest.

In figure 2.2, we see that the results we saw in the above regressions are generally consistent over the career of a player. The coefficients are comparable in both magnitude and statistical significance to the results in the tables and most have significance in of the 20-match bins. The figure suggests that the negative effects of early victories persist and appear to become stronger later in the career of a player. Note that this does not mean player performance falls over time: performance is increasing in match count and on net is increasing. However, as players play additional matches with each champion, the negative effect of good initial outcomes with that champion becomes stronger.

In addition to the primary results presented above, I run several robustness checks

[3] In fact, there are roles such as "support", where the player controls a champion whose role is to assist other players. Getting kills for oneself is sacrificed so that designated teammates gain the kills.

Table 2.3: Regression of Early Variables on Victory in Matches 50-70

| | Matches 50-70 | | | | Matches 10-30 |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| *Early performance and outcomes over first eight games* | | | | | |
| Early win rate | 1.19*** | -1.04** | -0.47 | -1.18*** | -2.10*** |
| | (0.33) | (0.41) | (0.30) | (0.41) | (0.21) |
| Kills | | 0.39*** | 0.49*** | 0.40*** | 0.65*** |
| | | (0.04) | (0.02) | (0.04) | (0.02) |
| Deaths | | -0.31*** | -0.19*** | -0.31*** | -0.35*** |
| | | (0.05) | (0.03) | (0.05) | (0.03) |
| Assists | | 0.05 | 0.04* | 0.05* | 0.10*** |
| | | (0.03) | (0.02) | (0.03) | (0.02) |
| *Win rate with other champions (over eight matches starting in indicated period)* | | | | | |
| After match 10 | | | | -0.97*** | |
| | | | | (0.33) | |
| After match 30 | | | | -2.06*** | |
| | | | | (0.33) | |
| After match 50 | | | | -3.36*** | |
| | | | | (0.33) | |
| Player Fixed Effects | Yes | Yes | No | Yes | Yes |
| N | 1,278,477 | 1,278,477 | 1,278,477 | 1,276,283 | 3,811,900 |
| $R^2$ Adj. | 0.004 | 0.004 | 0.003 | 0.004 | 0.006 |

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. Dependent variable is indicated in the column title. All regressions include player fixed effects as regressors. Standard errors, reported in parentheses, are clustered at player.

Table 2.4: Regression of Early Variables on Measures of Performance for Subsequent Matches 51 through 70

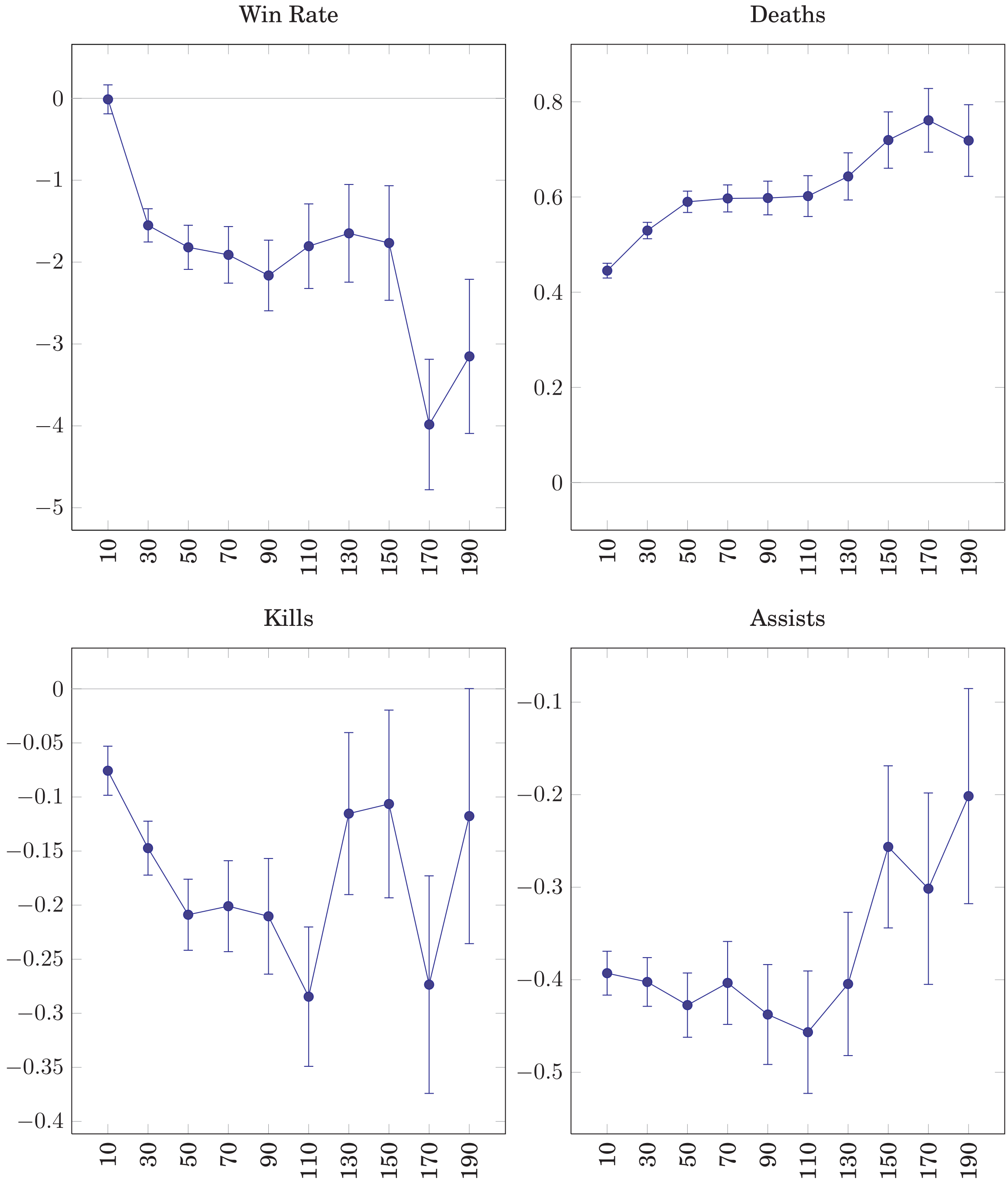
| | Simple Regression | | | With Controls | | |
|---|---|---|---|---|---|---|
| | (1) Kills | (2) Deaths | (3) Assists | (4) Kills | (5) Deaths | (6) Assists |
| *Early performance and outcomes over first eight games* | | | | | | |
| Early win rate | 0.07* | -0.08*** | -0.03 | -0.08* | 0.48*** | -0.32*** |
| | (0.04) | (0.03) | (0.04) | (0.05) | (0.03) | (0.05) |
| Kills | | | | 0.12*** | -0.02*** | -0.04*** |
| | | | | (0.00) | (0.00) | (0.00) |
| Deaths | | | | 0.01 | 0.14*** | -0.04*** |
| | | | | (0.01) | (0.00) | (0.01) |
| Assists | | | | -0.05*** | -0.03*** | 0.07*** |
| | | | | (0.00) | (0.00) | (0.00) |
| Player Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 1,278,477 | 1,278,477 | 1,278,477 | 1,278,477 | 1,278,477 | 1,278,477 |
| $R^2$ Adj. | 0.276 | 0.142 | 0.234 | 0.277 | 0.143 | 0.234 |

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. Dependent variable is indicated in the column title. All regressions include player fixed effects as regressors. Standard errors, reported in parentheses, are clustered at player.

Table 2.5: Robustness Check Related Regression of Early Variables on Victory

| | Matches 50-70 | | | | Matches 10-30 |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| *Early Performance and Outcome in First Eight Games* | | | | | |
| Early Win Rate | -1.04** | -1.06** | -1.20*** | -1.34** | -2.20*** |
| | (0.41) | (0.41) | (0.41) | (0.55) | (0.21) |
| Kills | 0.39*** | 0.39*** | 0.40*** | 0.40*** | 0.66*** |
| | (0.04) | (0.04) | (0.04) | (0.04) | (0.02) |
| Deaths | -0.31*** | -0.30*** | -0.31*** | -0.23*** | -0.35*** |
| | (0.05) | (0.05) | (0.05) | (0.06) | (0.03) |
| Assists | 0.05 | 0.05 | 0.05* | 0.05 | 0.10*** |
| | (0.03) | (0.03) | (0.03) | (0.04) | (0.02) |
| *Control (other Champion) Win Rate* | | | | | |
| First Eight Games | | -0.31 | -0.28 | -0.43 | -0.30* |
| | | (0.33) | (0.33) | (0.38) | (0.17) |
| *Control Win Rate in Other Periods* | | | | | |
| First Eight Games after Match 10 | | | -0.97*** | -0.72* | -2.11*** |
| | | | (0.33) | (0.37) | (0.17) |
| First Eight Games after Match 30 | | | -2.06*** | -1.98*** | |
| | | | (0.33) | (0.37) | |
| First Eight Games after Match 50 | | | -3.36*** | -3.40*** | |
| | | | (0.33) | (0.37) | |
| Player Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| N | 1,278,477 | 1,278,477 | 1,276,283 | 1,048,311 | 3,810,780 |
| $R^2$ Adj. | 0.004 | 0.004 | 0.004 | 0.004 | 0.006 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. Dependent variable is indicated in the column title. All regressions include mine fixed effects and quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. Standard errors, reported in parentheses, are clustered at the mine level.




Note: x-axis labels indicates start of 20 match bin. For example, “10” indicates matches 10-29 inclusive.

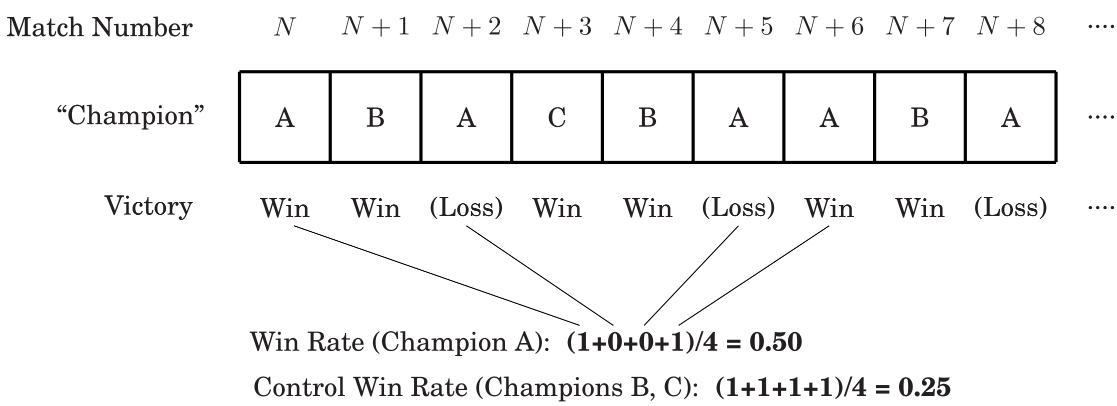
Figure 2.2: Plots of beta coefficients and standard errors of outcome variables over twenty match average starting at indicated match.

to examine and rule out the possibility of competing or confounding explanations for my results. I describe the results of these robustness checks below.

### 2.6.1 Robustness: Matchmaking System

In order to improve the quality of player experiences, League uses a matchmaking system that matches players by ability level, so that each team has approximately equal chances of winning. The matchmaking algorithm's evaluation of a player's abilities is influenced by the history of a player's success in previous games. Therefore, if a player does well in previous games they will be matched with some combination of harder opponents and weaker teammates. As a result, if a player is successful in past games for reasons other than their abilities, they will be more likely to lose in subsequent games. In this case, past success will be associated future defeats.

Figure 2.3: Illustration of the calculation of win rate and control win rate using a bin size of four




As a result, the matchmaking algorithm confounds the predicted effect of early win rate on future outcomes, as both predicted effects are negative. To address this, I evaluate the strength of this matchmaking effect by creating a control. I estimate the effect of the win rates of matches with *other champions* that occur around the same time of the early win rate.

Specifically, for each match observation where a player has played a given champion, I create an analogous statistic of win rate, similar to the early win rate of the champion played in the match observation, but for matches played with champions other than the champion of interest. Like the variable of interest which examines the win rate of each champion, this consists of the first eight games, starting at the time of the first play of the champion of interest. Figure 2.3 presents a diagram of how both the early win rate, and this control win rate is calculated. Note that since players have far from complete control over their choice of champions, this control win rate almost always overlaps the time period of the variable of interest. However, a portion of champions have the control period and later observed period overlapping. I drop these observations.

By constructing this contemporaneous win rate, I should be able to examine the magnitude of matchmaking responses on later win rate. I examine the effect on this control in Table 2.5. Columns 2-5 show the effect of this variable in various specifications. In particular, in column 2, introduces the control rate to the the preferred specification. The coefficient on this variable is less than a third of the magnitude of the main effect, and is statistically insignificant. Similarly, columns 3 and 4 show the control win rate being smaller and insignificant when controlling for other measures of player performance. While as expected the effect is negative, there is no evidence that matchmaking explains our main effect.

### 2.6.2 Robustness: Attrition Bias

Another possible reason for the negative coefficient on variable of interest is attrition bias. It is possible players who are more successful early on with a champion are more likely to play with that champion later. If large number of the later sample consisted of winners as a result, “regression to the mean” might result in a negative coefficient for mechanical reasons.

I find evidence for attrition. Table 2.6 shows simple regressions of the early win rate on the number of matches played with a champion by player. Columns 1 and 2 are regressions on the unprocessed dataset. Winning the early matches appears to have a strong effect on the number of matches played with a champion in the full sample, with a mean number of matches of approximately 10, the coefficient on the win rate ranges from approximately 2 to 6. Columns 3 and 4 in the same table uses the processed dataset. Here, notably, we see a small negative effect, which is difficult to explain. This may be related to the selected nature of the sample, which includes only players with a large number of games, as illustrated above in Figure 2.1.

Table 2.6: Regression of Early Outcomes on Number of Matches Played By Champion

| | Unprocessed Sample (Mean of Dep. Var. is 10.6) | | Process by Sample (Mean of Dep. Var. is 64.2) | |
|---|---|---|---|---|
| | (1) Number of Matches Played | (2) Number of Matches Played | (3) Number of Matches Played | (4) Number of Matches Played |
| *Performance in First 8 Games* | | | | |
| Early Win Rate | 6.45*** | 1.97*** | -2.45* | -9.40*** |
| | (0.04) | (0.05) | (1.38) | (1.75) |
| Kills | | 0.70*** | | 0.76*** |
| | | (0.01) | | (0.14) |
| Deaths | | -0.91*** | | -2.36*** |
| | | (0.01) | | (0.22) |
| Assists | | 0.07*** | | -0.38*** |
| | | (0.00) | | (0.13) |
| Player Fixed Effects | Yes | Yes | Yes | Yes |
| N | 2,748061 | 2,748,061 | 65,518 | 65,518 |
| $R^2$ Adj. | 0.12 | 0.13 | 0.08 | 0.08 |

* p < 0.1, ** p < 0.05, *** p < 0.01. Each column reports the results of one regression. Dependent variable is indicated in the column title. Standard errors, reported in parentheses, are clustered at the payer level.

While attrition appears to be negatively associated with early win rate in the processed dataset, I nevertheless design a specification to examine its influence. I run the preferred specification on a subsample of the data dropping observations where players win 6 or more of the 8 matches in the early period. In the subsequent sample, for the observations where frame of matches 31-50 exists, the early win rate in the full sample is 0.497 (N = 16502), and the win rate for the sample with these observations excluded is .453 (N = 13990). The trimming of high win rate observations should mechanically remove the effects of regression to the mean. The results of this regression is shown in column 4 of Table **??**. The coefficient of interest remains significant and is of the same magnitude.

### 2.6.3 Robustness: Complacency

There is another explanation for the main results: complacency. Instead of later performance being poorer as a result of an inefficient choice, the decline in performance may be the result of players working less in response to early successes, resulting in poorer later outcomes. The phenomenon may then reflect players preferences rather than the suboptimally of their decisions, and the efficiency implications are unclear.

I argue this is not the case. Firstly, the effect appears to persist for many matches afterward, representing months or even a year of gameplay. With so many matches in between, it seems implausible that players on the focus on success of matches that are all but ancient by the standards of online experiences. Instead, I argue that any effects of complacency should manifest much earlier. This seems more likely to result in other patterns of win rates, for example oscillating pattern of winning and losing where the average affect may be unclear, rather than the long, persistent effect we observe.

Secondly, while the early win rate has a negative association with later outcomes, other important measures of *individual* performance have strong positive associations with later outcomes. In Table 2.3 in all specifications, the number of player deaths predicts negative win rates, while kills, a positive measures of performance has significant positive coefficients. Similarly, in 2.4, other later outcomes such as "deaths" show the expected associations with early performance: more death' and kills predict more of the same in the future.[4] If complacency causes the robust negative relationship with early win rate and later outcomes, in contrast it does not explain why measures of performance are positively associated with good outcomes.

## 2.7 Conclusion

In this paper, I examine the strategic choice of champion that players repeatedly make in their matches for a competitive online game, League of Legends. I find evidence for the hypothesis that players are unduly influenced by chance of victory in their early matches with their champion.

[4]Note that other variables appear to have the opposite expected sign. For example, kills and deaths are positively associated. I argue this is explained with the roles players play: players with more kills play more "aggressive" roles involving direct combat, resulting both in more "kills" and "deaths". Similarly, "assists" are associated with a much less aggressive, supporting role, which is associated with fewer deaths and kills. There is also a confound in the fact that I cannot control for game length. Since longer game result in more of every statistic, there is likely a small but positive bias in all of the statistics in Table 2.4 .

The victory of a match is a stochastic variable because each match contains ten randomly matched players. Therefore on average each player must only have a minority of influence on each outcome. Without including controls for individual performance, I find an intuitive positive relationship between early and later match win rate, consistent with player skill being the underlying predictor in these simple regressions on success.

However, once controlling for individual player performance, we see that the outcome of early matches have an anomalous effect: early success leads leads to lower chances of victory later. This negative relationship also holds for for other measures of subsequent player performance. This is despite the fact that individual measures of early performance, where the player has much more direct influence is a positive predictor of subsequent performance. This counterintuitive result is explainable if players are making suboptimal choices, influenced by their early success. I also conduct a robustness checks which suggests my results are not due to statistical processes, or other idiosyncrasies in the game's matchmaking.

My paper is a examination of a multi armed bandit problem, which has no prior non-experimental precedents in the literature. While I chose my novel setting of an online game to advantage of the wealth of serendipitous data available, I argue that the problems that these players face in their competitive hobby are similar to many of the problems faced in more conventional economic settings. The low stakes decisions, over a long period of time, reflects many of the day-today choices faced in many common activities, where individuals must experiment and learn over time. In these settings, these random events on subsequent decisions may have an important and lasting impact.

## Appendix 2.A Basic Multi-Armed Bandit Model

The below formulation of the bandit problem exactly follows Bergemann and Valimaki (2006).

This is a discrete time infinite horizon Markov problem. At each period the decision maker chooses an action from a finite discrete set of choices $K$. $a_t \in \{1, ..., K\}$. If $a_t = k$, she gets the random payoff $x_t^k$ and the associated random variable is $X_k^t$. The state variable of the decision problem is given by $s_t$ and the distribution of $x_k^t$ is $F^k(\cdot; s_t)$. The state transition function $\phi$ is known to the decision maker, and depends on the choice and realized payoff $s_{t+1} = \phi(x_t^k; s_t)$.

As is common in all bandit problems, the objective is the discounted expected payoff with discount factor $\beta \in [0, 1)$ and the payoff from any choice $k$ only depends on outcomes when $a_t = k$.

Let $X^k(s_t^k)$ be the random variable with distribution $F^k(\cdot; s_t^k)$. The decision maker's problem is to maximize the value function:

$$V\left(s_0\right) = \sup_a \left\{ E \sum_{t=0}^{\infty} \beta^t X^{a_t}\left(s_t^{a_t}\right) \right\}$$

The major result is from Gittins and Jones (1974) who show that finding the optimal policy can be reduced to solving $K$ one-dimensional problems. For each choice $k$ one needs to calculate the following index, which only depends on the state variable of index $k$. The index takes the form

$$m^k\left(s_t^k\right) = \sup_\tau \left\{ \frac{E \sum_{u=t}^{\tau} \beta^t X^k\left(s_u^k\right)}{E \sum_{u=t}^{\tau} \beta^t} \right\} \tag{2.2}$$

Then the Gittins index theorem states the optimal choice in each period is to select the choice with the highest index.

**Theorem 2.A.1** (Gittins and Jones (1974))**.** *The optimal policy satisfies $a_t = k$ for some $k$ such that*

$$m^k(s_t^k) \geq m^j(s_t^j) \text{ for all } j \in \{1, ..., K\} \tag{2.3}$$

Note that $m^k$ in 2.2 is increasing in $X^k(s_t^k)$. If the agent gets a draw $s_{t+1}$ from $\phi$ that produces a higher $F^k(\cdot; s^k_{t+1})$, the agent is more likely to choose $k$.

Intuition for this result is provided in Weber et al. (1992) and Bergemann and Valimaki (2006).

# Chapter 3

# Productivity, Safety, and Regulation in Coal Mining: Evidence from Disasters and Fatalities[1]

## 3.1 Introduction

Since 1900, over 100,000 workers have been killed in coal mines in the United States (Alford, 1980). Coal miners are exposed to a wide range of hazards including gas explosions, shifting rock, falls, and machinery and automotive accidents. Yet, as a result of government regulations, technological change, and a general emphasis on safety, coal mining fatalities have dropped dramatically over the twentieth century. Today, in any given mine, fatalities are now an unexpected and extraordinary event. Though fewer than before, coal mining deaths still occur in the United States: between 2000 and 2014, there have been 446 fatalities and five disasters – events with five or more deaths.[2] Injuries of all kinds are still commonplace in this sector.

The fact that safety remains an important concern in the coal mining industry suggests that it is appropriate to think of the industry as producing two joint outputs: extracted mineral and safety. The purpose of this paper is to estimate the joint production function for mineral and safety in the coal mining industry and to

[1]Coauthored with Gautam Gowrisankaran, Eric Lutz, and Jefferey Burrows. An equal amount of work was done by the authors.

[2]Source: authors' calculation from MSHA data, ending at Q3:2014.

evaluate the role of regulation in affecting this tradeoff. Understanding the relationship between coal output, safety, and regulation is important for evaluating how productivity and safety might continue to increase in the coal-mining industry. The coal-mining industry employs over 90,000 workers in the United States (Energy Information Administration, 2014), underscoring the importance of this goal. In addition, understanding the tradeoff between productivity and safety might inform us about the potential sources of productivity increases in commodity extraction and manufacturing industries more generally, many of which are as or more dangerous than coal mining.[3] Finally, we are interested in how regulation affects safety and productivity. What is the marginal value of increased MSHA enforcement activity? Do additional inspections and enforcement of regulations compromise productivity in the name of safety, enhance both safety and productivity, or simply have no effect on safety?

To analyze our research question, we make use of a panel dataset that records location, productivity, accidents, and regulatory inspections for coal mines throughout the U.S. The detailed data allow us to measure output much more precisely than in most industries, and to pinpoint potential factors that influence productivity, including the role of regulatory activity. Despite the uniquely detailed data, it would be problematic to answer our research question with a regression of a proxy of safety (such as accidents) on productivity. Unobserved factors, such as management quality or ease of coal extraction at a mine, might impact both the observed safety level and productivity. For instance, firms with higher management quality may achieve higher levels of both safety and productivity. Certain mines, such as surface mines, may allow for greater ease of coal extraction and also be inherently less dangerous than underground mines. Such factors would then result in an endogeneity bias.

We employ a different identification approach that does not rely on input choices being orthogonal to unobserved management quality. Instead, we identify the joint production function of coal output and safety by using disasters and fatalities as a source of quasi-experimental variation that affects the relative price of safety and productivity by increasing the cost of future accidents. Our identification strategy relies on the presumption that fatalities and disasters are tail events that are

[3] Although coal mining was once one of the most dangerous occupations by fatality rate, it is no longer in the top 10. Occupations with higher or similar death rates include fishers (117.0 per 100,000 full time yearly worker equivalents), logging workers (127.8), roofers (40.5), and cement and concrete manufacturing (16.0). The rate for coal mining is 18.0. (Source: U.S. Bureau of Labor Statistics. *http://www.bls.gov/iif/oshwc/cfoi/cfoi_rates_2012hb.pdf*.) The large number of deaths in the coal mining industry reflects the large number of workers in the industry.

generally unexpected; certainly systems exist to reduce severe risks and no fatality would be allowed to occur if it could be foreseen by workers or management.

Why might a fatality or disaster change the relative price of safety to mineral output? We hypothesize that a fatality at a mine might increase the cost of future accidents through a variety of mechanisms. Most directly, a fatality at a mine might increase MSHA regulatory inspections. This might then make the mine more likely to suffer penalties from safety violations, which would increase the price of safety relative to mineral output. Although regulation is one way in which fatalities might change the relative price of safety relative to mineral output, there are many other mechanisms. For instance, mines may become stigmatized by fatalities, which may harm their ability to attract labor or investment. Experiencing a fatality may also increase the psychic costs of risk exposure to both workers and management through fear or guilt. Finally, firms may have incomplete information about the safety levels of their mines and a fatality may be a signal that the mine is more dangerous than previously believed, implying that improving safety becomes a better investment following a fatality.

In contrast to single fatalities, mine disasters are generally followed by intense media exposure and public reprobation of those responsible for the disaster. For example, in the aftermath of the 2010 Upper Big Branch Mine disaster in Raleigh County, West Virginia, in a public eulogy to the fallen miners, President Obama remarked, "owners responsible for conditions in the Upper Big Branch Mine should be held accountable for decisions they made and preventive measures they failed to take" (Obama, 2010).

The publicity after the disaster was sufficient for Nike, Inc. to air a national television ad referencing the fallen miners (which was later pulled due to controversy).[4] Given the media exposure, we hypothesize that, relative to single fatalities, disasters are more likely to affect the cost of future accidents in the broader geographic area than just at the particular mine.

To formalize the production function of coal output and safety, and the effect of disasters and fatalities on this production function, we first develop a simple neoclassical model of the production of safety and mineral output. In our model, firms choose labor and safety inputs that, together with random draws, lead to a production level production for mineral output and accidents. The expected number

[4] CBS (2010).

of accidents per hour and mineral output are both monotonically decreasing functions of the chosen safety level implying that safety and mineral output are substitutes. Thus, this model hypothesizes a tradeoff between productivity and accident risk.

Within this simple economic framework, firms choose a level of safety that balances the marginal reduction in accident cost against the marginal reduction in productivity from an additional safety unit. We then assume that disasters and fatalities are very low probability shocks that increase the cost of future accidents. We show that given intuitively reasonable conditions on the interactions between labor and safety inputs and between safety input and the cost of accidents, an increased cost of accidents leads to greater safety input, fewer workers, and hence less mineral production. Thus, the model predicts that mines react to a disaster or fatality by increasing safety at the cost of less mineral output per worker and fewer workers. These testable predictions form the basis for our empirical work: we examine whether mines reduce accident rates after a disaster or fatality, and if there exists an associated cost to productivity. We also directly test one set of mechanisms by which the cost of accidents might increase, by examining whether MSHA enforcement activity increases after a disaster or fatality.

The model above assumes that mining companies behave rationally. We imagine firms walking a tightrope, attempting to maximize production while sensibly minimizing expected accidents given their production level. But in reality is this how firms actually behave? Do firms optimally choose mineral output and safety, or could other circumstances affect output? Surprising and robust evidence suggests that some industries have unrealized productivity. In his paper on iron ore producers, Schmitz (2005) finds that labor productivity doubled through changes in work practices. Foreign competition was the shock that spurred these new practices in the iron ore industry.[5] Similarly, it is possible in the aftermath of a fatality or disaster shock that an increased focus on safety could also spur productivity improvements in the coal mining sector. Supporting this view, occupational health and safety research has found that both productivity and safety could increase from managerial attention and training in underground mining (Fiedler et al., 1984), logging (Montorselli et al., 2010), and construction (Everett and Slocum, 1993). By testing the sign and magnitude of the impact of fatalities and disasters on productivity and safety, our empirical analysis provides a test of our simple

[5] Hendel and Spiegel (2014) also find large and not easily explainable productivity changes in Israeli steel mini-mills.

neoclassical production function model.

Our empirical analysis proceeds as follows: we create a panel at the mine-quarter level that merges several publicly available MSHA datasets, including the *Accidents Injuries Dataset*, the *Employment/Production Data Set (Quarterly)*, the *Inspections Data Set*, and the *Violations Data Set*. Key variables in our data include mine location, coal production, hours worked, the number of fatalities and other accidents, and information on MSHA inspections and citations. Our regressions all use an event analysis framework. We regress dependent variables – such as productivity, accident rates, and MSHA inspections – on the occurrence of a fatality in the same mine or a disaster in the same state within the two previous years. In all regressions, we use mine and time fixed effects, and we cluster at the mine level. As a falsification exercise, all specifications also examine future fatalities or disasters within the two subsequent years. We also drop the five mines with a disaster from our sample.

We first examine "first stage" results of the impact of lagged fatalities at a mine on MSHA inspections. We find that MSHA inspections significantly increase during the quarter of the fatality, and that this increase is sustained for two years. The magnitudes are large, with a 11% increased inspection rate two years after the fatality. In contrast, the inspection hours and penalties only significantly increase for the first year after the fatality. Together, these results suggests that mines react to the increased MSHA inspections in the first year after a fatality by resolving problems that could lead to citations. Hence, the inspections become shorter and ultimately, the inspectors find no more to penalize than in the baseline. In almost all our regressions, we find that future fatalities had no statistically significant effect, suggesting the absence of pre-existing trends.

Having established at least one causal pathway by which fatalities might increase the cost of accidents, we turn to understanding the impact of fatalities on productivity and accidents. Here we find no evidence that fatalities decrease severe accident rates (defined as the rate of fatalities and permanent disabilities per hours worked) but find that they decrease less-severe accident rates (defined as the rate of all other accidents) by 10% two years after a fatality. There is also no evidence that productivity increased. Thus, the results indicate that the increased inspections and other effects from a fatality only affect less-severe accidents, but that they do this with no apparent negative effect on productivity.

We next consider disasters, and examine the effect of having a mine in the same state experience a disaster (omitting the mine with the disaster itself). Here, we find no evidence of increased MSHA regulatory scrutiny at other mines within the same state. However, we find that accidents drop significantly and by a large amount following a disaster. In particular, the rate of less-severe accidents per hours worked decreases by 23% and the rate of fatalities decreases by 68% two years after a disaster. We also find potential evidence that productivity went down, by 7% following a disaster, a result that is marginally significant (P=0.099). However, the pre-trends on productivity, while not statistically significant, are similar in magnitude, limiting the plausibility of any finding of a productivity decrease. Nonetheless, this evidence is supported by regressions using state-year-level data that the number of managers and supervisors at mines increases 11% two years after a disaster in a state, with no significant change in other workers.

It is useful to understand the cost savings that mines may incur from the drops in accident risk. An influential review article, Viscusi and Aldy (2003), finds the value of statistical life lies between \$4 and \$9 million using U.S. labor market estimates (which is close to our context). Using the midpoint value of \$6.5 million (Viscusi and Aldy, 2003), we find that the reduction in risk of fatalities is worth \$1.41 per hour worked. For less-severe accidents, a \$30,000 estimate (National Safety Council, 2014) implies a cost savings of \$0.24 per hour worked from the reduction in this type of accident. Hence, we believe that the total dollar cost savings to the firm from the decreased accidents following a mine disaster may be in the range of \$1-\$2 per hour worked. As a point of comparison, the 7% drop in productivity noted above, if real, would at the least imply a need to add 8% extra work hours to mine the same coal. At \$25/hour, this represents an extra \$2 in wages per current hour worked. However, labor costs are only a small part of the total costs of coal extraction. If lost production from enhanced safety could not be recouped with just labor input, the reduction in coal output would cost up to \$14 per hour worked. Thus, the value of potential productivity losses here is not trivial.

The remainder of the paper is organized as follows: Section 2 provides background information on the industry and literature. Section 3 exposits our model and estimating framework. Section 4 discusses the data. Section 5 discuss results. Finally, Section 6 concludes.

## 3.2 Background

Coal mining is an important industry to study to answer questions regarding safety, productivity, and regulation. Tens of thousands of workers have been killed in the United States in the history of coal mining and many more have been injured and disabled. Figure 3.1 shows total annual U.S. mining deaths from 1900-93. During this long period, mining deaths and morbidity in the U.S. declined steadily. However, mining remains a dangerous occupation, with disasters such as in the Sago Mine (2006) and in Upper Big Branch (2010) killing 12 and 29 workers, respectively.

In the late nineteenth and early twentieth century, each miner worked in a "room," which is a small area of the mine that is individually allocated to a particular miner. A frequent cause of fatalities was a roof collapse in the miner's room (Fishback, 1992). The room's roof was progressively weakened by the process of coal extraction. Therefore, each miner had control over his or her safety, as miners spent their days literally demolishing the columns supporting the roofs over their heads. Pressure on the remaining coal increased as coal was removed, which actually made further mining easier (suggesting the opposite sign to the endogeneity between mineral output and safety noted in the introduction). As miners were paid piecerate, this "softening" was one reason miners valued obtaining the maximum possible coal from their room, even under conditions most would find terrifying.[6] Skilled miners could reasonably estimate when a roof was about to collapse, and dig the furthest, but it was never possible to avoid collapse with certainty. Each miner worked with the knowledge there was some low but real possibility that the roof would collapse, and some were killed when this in fact occurred.

These roof collapses are a vivid and unusually direct example of balancing productivity with safety. Although this hazard no longer exists, tradeoffs can be found virtually anywhere there is a risk of injury in a mine. For example, speed limits on trucks or other machinery can improve safety but slow production, and safety equipment such as gloves protect workers but reduce dexterity. Mine construction involves tradeoffs between speed and structural considerations; miners still suffer fatalities from mine collapses. Finally, allowing for extra escape routes, rescue areas, and training reduces the time spent on coal extraction, but may reduce deaths in the events of a collapse.

[6] A vocabulary developed that described the various sounds the roof could make. The sound of a roof groaning under reduced support was known as the "roof working". Some sounds resembled crashing thunder. It is claimed experienced miners could detect a distinct sound that indicated imminent collapse (Brophy and Hall, 1964).

In response to the dangers from mines, significant safety regulations have been enacted at the U.S. federal government level. The main regulatory changes are the creation of the Bureau of Mines in 1910, the Coal Mine Safety and Health Act of 1969,[7] the Occupational Safety and Health Act (OSHA) of 1971 (Alford, 1980), the Federal Mine Safety and Health Act of 1977 (which created the Mine Safety and Health Administration, MSHA) (Weeks and Fox, 1983), and, most recently, the Mine Improvement and New Emergency Response (MINER) Act of 2006. The MINER Act further advanced rules, outlined in the 1977 Act, pertaining to emergency response, emergency evaluation and notification, post-accident communications and tracking, mine rescue teams and their equipment, and sealing of abandoned areas of underground coal mines.

In many cases, these regulations were spurred by mine disasters. For instance, the 1952 act followed the 1951 Orient #1 explosion that killed over 100 miners, the 1969 act followed the 1968 Consol #9 disaster that resulted in 78 coal miner deaths, and the 1977 act followed the 1972 Sunshine Mine fire that killed 91 miners. Today, coal mines operate in a strict liability operating environment with mandated MSHA inspections – a minimum of 2 times per year for all surface mines and 4 times per year for all underground mines. Following inspections, the government can enforce regulations with punitive damages for statute violations that can range up to hundreds of thousands of dollars per violation.

Figure 3.1: Historical Coal Mining Fatalities in the United States.

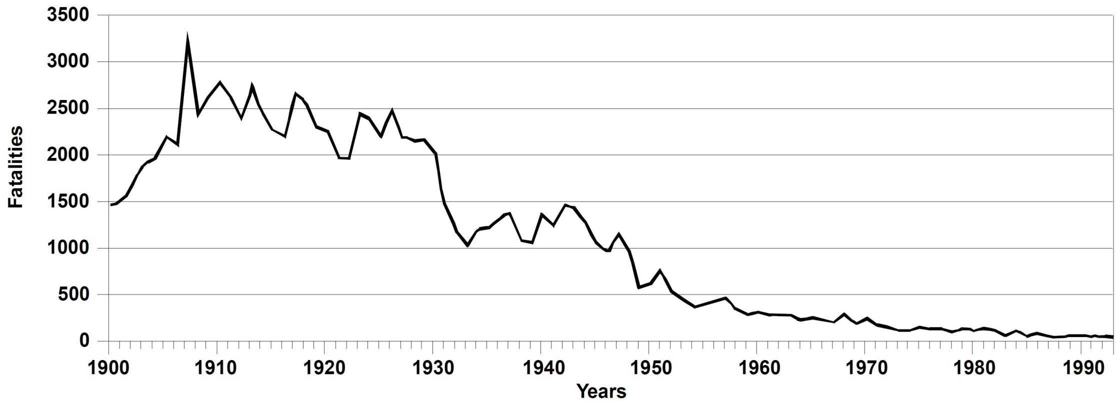



Figure taken from *Energy Information Administration. Coal Data: A Reference*. DIANE Publishing Co, 1995. Page 71 (Figure 27).

A small literature has formalized the tradeoffs between safety, productivity, and regulation in coal mining and other sectors. To our knowledge, the first to do this for

[7]See http://www.msha.gov/REGS/ACT/ACTTC.HTM.

coal mining is Sider (1983), who specifies a model of coal mining with tradeoffs between safety and production, on which our model builds. Sider (1983) estimates a Cobb-Douglas production function for coal mines in Illinois that is motivated by his model. He did not estimate the reduction in safety stemming from regulations nor did he attempt to control for the endogeneity of safety and production choices. Gray (1987), who considers all manufacturing sectors, estimates a production function that includes measures of regulations. He finds that safety regulations – such as OSHA inspections – contributed as much as 30% of the decline in productivity growth during the 1970s. Kniesner and Leeth (2004) examine whether MSHA enforcement activities reduce mine injuries and find very small effects.

At the same time, an empirical literature finds that productivity and safety may be improved simultaneously in a number of different sectors. Hausman (2014) finds that electricity market restructuring allowed nuclear power plants to operate both more safely and more efficiently. In a detailed evaluation of four Italian logging crews, Montorselli et al. (2010) find that the crew with formal safety training had both the best safety record and the highest productivity. Fiedler et al. (1984) find that mine management safety programs simultaneously improve productivity and reduce accidents.

Finally, another literature seeks to understand what and who determines safety choices. Sawacha et al. (1999) analyze safety at construction sites, finding that incentive or bonus pay can lead to decreased safety. Similarly, Hensher et al. (1992) studied long-haul truck drivers and found that drivers respond to financial incentives by reducing safety. Both of these studies find that workers shift priorities when incentives change, thereby demonstrating that workers may have some control over the tradeoff between safety and productivity. Other studies show that management may also have control over safety and may respond to incentives concerning worker safety (Rittenberg and Manuel Jr, 1987).

We believe that our study contributes to the general understanding of the determinants of productivity and safety through our novel identification strategy of using the random variation in mine disasters and mine fatalities. This allows us to estimate the tradeoff between productivity and safety in a way that controls for the endogeneity of safety and labor input choices.

## 3.3 Model and Estimation

### 3.3.1 Model

We present a simple neoclassical model of safety and productivity in mine operation and use the model to develop testable implications. In our model, profits are determined by the revenues from coal sales, the cost of the labor input, and the cost of accidents. The mine chooses a labor input of hours worked, $x$, and a safety input, $s$. Together, these choices lead to a stochastic occurrence of accidents and coal production. The mine is faced with – but does not choose – a cost of accidents, $d$, which is determined by the past occurrence of fatalities and disasters. A high cost of accidents implies that a given accident will cost the firm more.

We now turn to the details of the model. Let $x$ and $s$ both be non-negative and assume that a higher value indicates more hours worked or a higher safety input, respectively. Let the random vector $A(s)$ denote the occurrence of accidents per hour worked. We allow $A(s)$ to be vector valued, to account for multiple types of accidents. We expect that $s = 0$ would lead to a high expected number of accidents per hour and that a large $s$ would lead to very few accidents. Denote the cost of accidents per hour as $C(A(s), d)$. Finally, we will want to directly consider the cost of any safety input and accident price $d$ rather than the occurrence of accidents. Let the random variable $c(s, d) \equiv C(A(s), d)$.

Using this notation, we write the expected profit function for a mine as:

$$E[\pi(x, s|d)] = E[F\,(x, s) - wx - c\,(s, d)\,x]. \tag{3.1}$$

In (3.1), the first term, $F(x, s)$, is the net coal production expressed in dollars. In the second term, $w$ is the wage and $wx$ is the total wage bill. More generally, we might think of $x$ as all factor inputs and $wx$ as the costs for all factor inputs. Last, $c\,(s, d)\,x$ is the total accident cost, which is the per-unit cost of accidents multiplied by the number of hours worked.

We now specify assumptions regarding the impact of the inputs on profits in (3.1).

*Assumption* 3. Impact of inputs on expected profits

1. Expected coal production $E[F(x, s)]$ is increasing in $x$ and decreasing in $s$;

2. Expected accident cost $E[c(s,d)]$ is decreasing in $s$ and increasing in $d$.

We believe that Assumption 3 is intuitive. Part 1 of Assumption 3 implies that firms produce more coal if they employ more hours (conditioning on the safety input), and less coal if they use more safety input (conditioning on hours worked). Part 2 implies that expected per-hour accident cost is decreasing in the safety input $s$ and increasing in the cost of accidents, $d$.

We next make assumptions on the second derivatives that we also believe are intuitive.

*Assumption* 4. Impact of second derivatives of inputs on outputs

1. All functions are twice differentiable;
2. $\partial^2 EF(x,s)/\partial x \partial s$ is negative;
3. $\partial^2 Ec(s,d)/\partial s \partial d$ is negative.

Part 1 of Assumption 4 is just for simplicity; we could alternately derive our results using lattices. Part 2 states that the marginal product of labor, in terms of mineral output, decreases in the chosen level of safety input. This result would be generated if, for instance, $s$ were monotonic in the fraction of the time that workers spent on safety training, and workers allocated their time between safety training and coal extraction. In this case, with a higher $s$, an additional hour of labor would result in a lower fraction of an hour spent on coal production and hence in less coal. Part 3 implies that the marginal cost from increased accident risk (or equivalently, marginal benefit from increased safety input) is increasing as we increase the cost of accidents $d$. This would be generated if the cost of accidents were linear in $d$ for instance.

Since we assume differentiability, we can write the first-order condition with respect to labor input as:

$$\frac{\partial E[\pi(x,s|d)]}{\partial x} = 0 \Rightarrow \frac{\partial E[F(x,s)]}{\partial x} - w - E[c(s,d)] = 0, \tag{3.2}$$

and with respect to safety as:

$$\frac{\partial E[\pi(x,s|d)]}{\partial s} = 0 \Rightarrow \frac{\partial E[F(x,s)]}{\partial s} - x\frac{\partial E[c(s,d)]}{\partial s} = 0. \tag{3.3}$$

Equation (3.2) states that firms set their marginal product of labor input with respect to mineral production equal to the wage plus the extra accident cost. It differs from a standard production FOC in the third term, which is the extra accident cost from the additional hour of work. Equation (3.3) states that firms set their safety input so that the cost of increasing safety at the margin in terms of reduced expected mineral output is equal to the benefit of safety at the margin in terms of decreased total expected accident costs.

We now turn to our main result, which is that we will observe a decrease in mineral output and in expected accidents per hour (or equivalently, an increase in safety input) following a disaster or fatality. We define $x^*(d)$ and $s^*(d)$ to be the profit maximizing choices of labor and safety inputs respectively. Formally:

*Proposition* 2. $x^*(d)$ is decreasing in $d$ and $s^*(d)$ is increasing in $d$.

*Proof.* Assumptions 3 and 4 imply that:

$$\frac{\partial^2 \pi}{\partial x \partial d} = -\frac{\partial E[c(s,d)]}{\partial d} < 0,$$

$$\frac{\partial^2 \pi}{\partial x \partial s} = \frac{\partial^2 E[F(x,s)]}{\partial x \partial s} - \frac{\partial E[c(s,d)]}{\partial s} < 0,$$

and

$$\frac{\partial^2 \pi}{\partial s \partial d} = -x \frac{\partial^2 E[c(s,d)]}{\partial s \partial d} > 0.$$

Amir (2005) provides a simple proof of the monotonicity of optimal choices that is based on supermodularity as developed by Topkis (1978). For convenience, define $y = -x$ and $y^*(d) = -x^*(d)$. Then, the first condition of Amir (2005) Theorem 9 requires that the second derivatives of $\pi$ with respect to $y$, $s$, and $d$ all be positive, which we have shown. The second condition of the theorem is satisfied by our assumption that the choice set for $x$ and $s$ includes all non-negative real numbers, irrespective of $a$. Thus, by Amir Theorem 9, $s^*(d)$ and $y^*(d)$ are increasing in $d$, implying also that $x^*(d)$ is decreasing in $d$. ■ □

We believe that this result is intuitive. In a neoclassical model, firms choose safety levels to balance the accident cost with the production losses from safety. An unexpected fatality or disaster increases the cost of future accidents. This increases the firm's marginal benefits to improving safety, causing the firm to choose a greater safety input. The marginal productivity of labor in producing mineral output falls as a result, causing the firm to chooses less labor input. Together these effects result in a decrease in mineral output.

### 3.3.2 Estimation framework

Our model predicts that a fatality or disaster shock will lead to an increase in the optimal safety input $s^*$ following a disaster or fatality. An increase in safety input in turn implies a decrease in the expected accidents per hours worked, $A(s^*)$. It also implies a decrease in the mineral output per worker, or productivity, conditioning on the number of hours worked. Finally, it implies a decrease in the number of hours worked.

Thus, we specify regressions that examine whether accidents per hour, productivity, and hours worked are affected by past mine disasters or fatalities. Specifically, we perform regressions at the mine-quarter level. Let $i$ denote a mine and $t$ denote a calendar quarter. We perform regressions of the form:

$$A_{it} = \alpha_i + \gamma_t + d_{it} + X_{it}\beta + \varepsilon_{it}, \tag{3.4}$$

where $A_{it}$ are accident rates per hours worked, $\alpha_i$ are mine fixed effects, $\gamma_t$ are time (quarterly) fixed effects, $d_{it}$ is the current cost of accidents (as determined by lagged fatalities and disasters), $X_{it}$ represent other covariates such as hours worked, and $\varepsilon_{it}$ include shocks to accidents rates. Besides accidents, we also perform similar regressions to (3.4) but using productivity and hours worked.

We allow for two types of regressions, based on $d_{it}$ indicating either disasters or fatalities. For our regressions where $d_{it}$ indicates disasters, we specify that the price of safety may change if there is a disaster located near the mine. The reason for this is because, as noted above, disasters are widely known and have intense media and political scrutiny. For our regressions where $d_{it}$ indicates a fatality, we do not believe that there will be a wider change in the price of safety. Hence, we specify that the price of safety is likely to change only for the mine with the fatality itself. In all cases, we drop the five mines with a disaster, because other factors besides a change in the price of safety, such as physical damage, may be affecting their mineral output and safety choices. Noting that $A_{it}$ is a vector, we estimate different versions of (3.4) with different types of accidents.

In (3.4), we include mine fixed effects because each mine produces with a different technology. For instance, we would expect underground mines to have more risks than surface mines, all else equal. By including mine fixed effects, we are controlling for the baseline risk at each mine in evaluating how past disasters and fatalities affect safety choices. Our estimates thus reflect the change in accidents after the disaster or fatality. Similarly, our covariates include time fixed effects because safety technologies and regulations have changed over time and because different input prices might cause mines to make different tradeoffs over time.

In addition to the regressions based on (3.4) we also specify first stage regressions that consider the mechanisms whereby a past disaster or fatality might increase the cost of future accidents. The central mechanism that we consider is regulation through MSHA. The format of these regressions is the same as (3.4) except that the dependent variable here indicates MSHA regulatory visits and enforcement.

Our regressions all include *future* fatalities or disasters. This inclusion forms a falsification test that would allow us to reject the causal interpretation of a disaster or fatality. For instance, if we observe that a future fatality significantly predicts higher accident risk, we might infer that the mine has undergone a period of low safety relative to its long-run average. We might then believe that the higher accident risk led to the fatality, rather than that the fatality changed the price of safety which led to differential accident risk, as we assume in our model. Thus, our model will tend to be more plausible to the extent that the future indicators are not significant predictors.

We cluster all our standard errors at the mine level to allow for $\varepsilon_{it}$ to be serially correlated over time. We also weight all regressions by the mean hours worked at a mine, because mines with more hours worked may provide more information.

Finally, we include some regressions where the unit of observation is the state-year. These regressions use data on occupations within the mining sector that are only available at this level of aggregation. For these regressions, we cluster standard errors at the state level and weight regressions by the mean number of workers in the state.

## 3.4 Data

We obtain most of our data from the Mine Safety Health Administration (MSHA). We merge several datasets, all of which are available from the “Data Sets” area on the MSHA.gov website. The datasets report information on coal and metal mines. We keep from them exclusively records pertaining to coal mines. The *Employment/Production Data Set (Quarterly)* indicates the total coal production, the number of employees, U.S. state of location, and number of hours worked for all coal mines in the U.S. at the quarterly level from 2000 through the third quarter of 2014.

The *Accident Injuries Data Set* reports detailed information for all coal mining accidents in the U.S. From this database, we use the reported degree of injury. A fatality is a degree 1 injury, and an injury resulting in a permanent partial or total disability is a degree 2 injury. We classify accidents that result in degree 1 and 2 injuries as severe accidents. We classify injuries of degrees 3-6 as less-severe accidents. We exclude injuries of degrees 7-10, which include injuries from natural causes and injuries from non-employees.

The *Violations Data Set* reports the number of violations and financial penalties mines were assessed from MSHA inspectors during each quarter. The *Inspections Data Set* reports the number of inspections in a mine and the total hours each mine is inspected. Finally, the *Mines Data Set* provides latitude and longitude information for each mine.

Table 3.1 provides details on injury by degree during the period of our data. Fatalities are the most rare, with 446 observed over this period. Permanent disabilities occur slightly more often than fatalities, with 785 observed over the same period. Among the less-severe accidents, the most common are accidents with injuries that require days away from work only, with 42,122 observed, followed by ones that require no days away from work, but still require medical treatment, with 21,211 observed.

Figure 3.2 graphs fatalities and disasters over time during the period of our data.[8] We define a disaster as an accident with five or more fatalities. In this time period, all disasters are caused by the ignition of explosive gasses. We observe five disasters, two in the state of West Virginia, and one in each of Utah, Kentucky and Alabama.

[8] We exclude 2014 from this figure as have only 3 quarters of data from 2014.

Table 3.1: Summary of Accident Occurrences in Sample

| Injury Degree | Accident Description | Severe Injury | Number Observed |
|---|---|---|---|
| 1 | Cases resulting in death | Yes | 446 |
| 2 | Cases with permanent total or partial disability | Yes | 785 |
| 3 | Cases with days away from work only | No | 42,122 |
| 4 | Cases with days away from work and restricted work | No | 3,854 |
| 5 | Cases with days of restricted work only | No | 4,609 |
| 6 | Cases without days away from work (but with medical treatment) | No | 21,211 |

Note: sample period of accidents is is Q1:2000 through Q3:2014.

All disasters, including the mine and state in which they occurred, are noted on the figure.

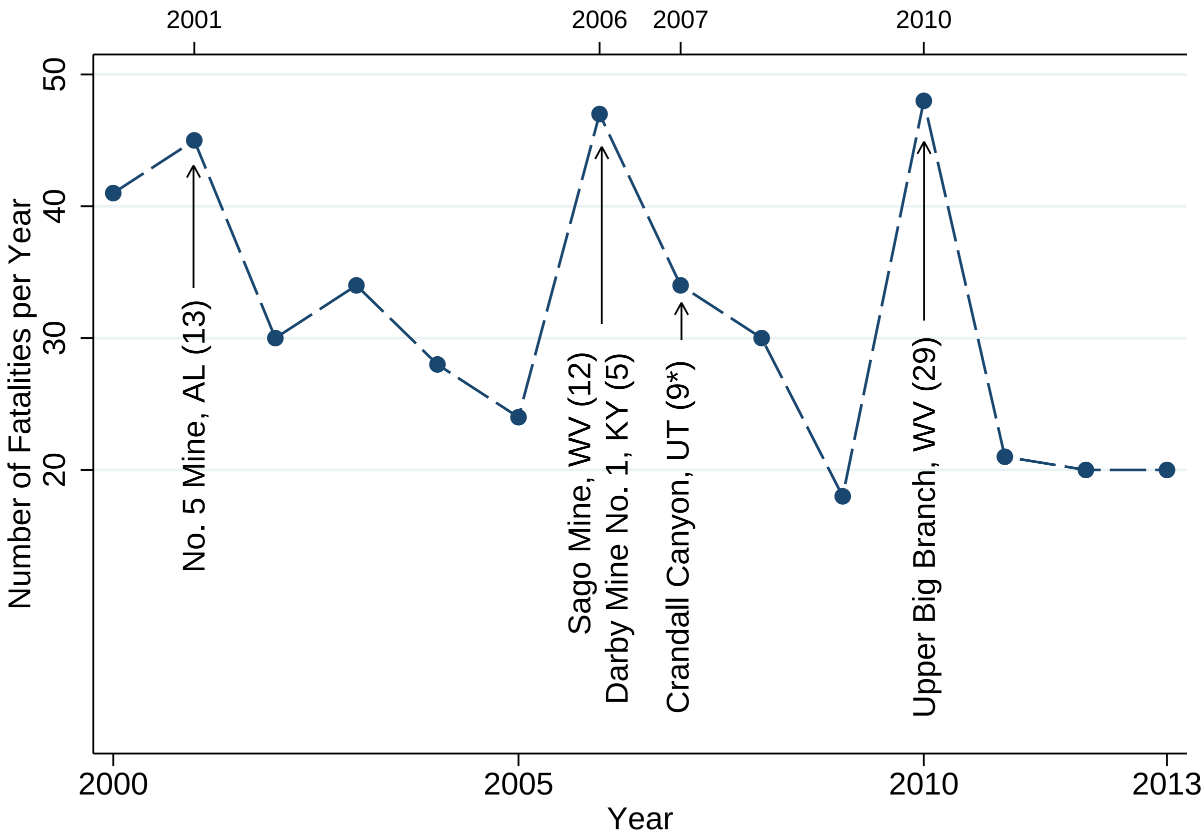
We chose five as the threshold number of fatalities that defines a disaster as we thought that this number would represent a cutoff for generating more widespread attention. Note also from Figure 3.2 that four out of the five disasters that we observe have far more than five fatalities. Also, most non-disaster fatalities represent single fatalities. In particular, we observe 322 mine/months with 1 fatality, 25 mine/months with 2 fatalities, 2 mine/months with 3 fatalities, and none with 4 fatalities. We believe that this evidence shows that there is a sharp division between the disasters and the non-disaster fatalities.

Figure 3.3 graphs severe and less-severe accidents over time during the period of our data. We observe a downward trend in less-severe accidents over time. There is no clear trend for severe accidents.

The production and accident databases are reported separately by each subunit within a mine. To create our estimation sample, we remove all observations that are in subunits that indicate office work instead of mining activity. We then collapse the data so that our unit of observation is a mine observed over a quarter. We define productivity as coal production measured in tons divided by person-hours of labor. Our data on MSHA violations and actions are reported at the mine-quarter level so we do not collapse these data.

Our analysis sample drops mines which produced no coal ever during our sample period. We also drop mine-quarter observations which reported fewer than 2,000

Figure 3.2: Fatalities and Disasters in Coal Mines




Note: number of deaths for each disaster reported in parentheses. *Of the nine fatalities associated with the Crandall Canyon disaster, three were rescue workers who were killed ten days subsequent to the initial collapse.

Figure 3.3: Severe and Less-Severe Accidents in Coal Mines

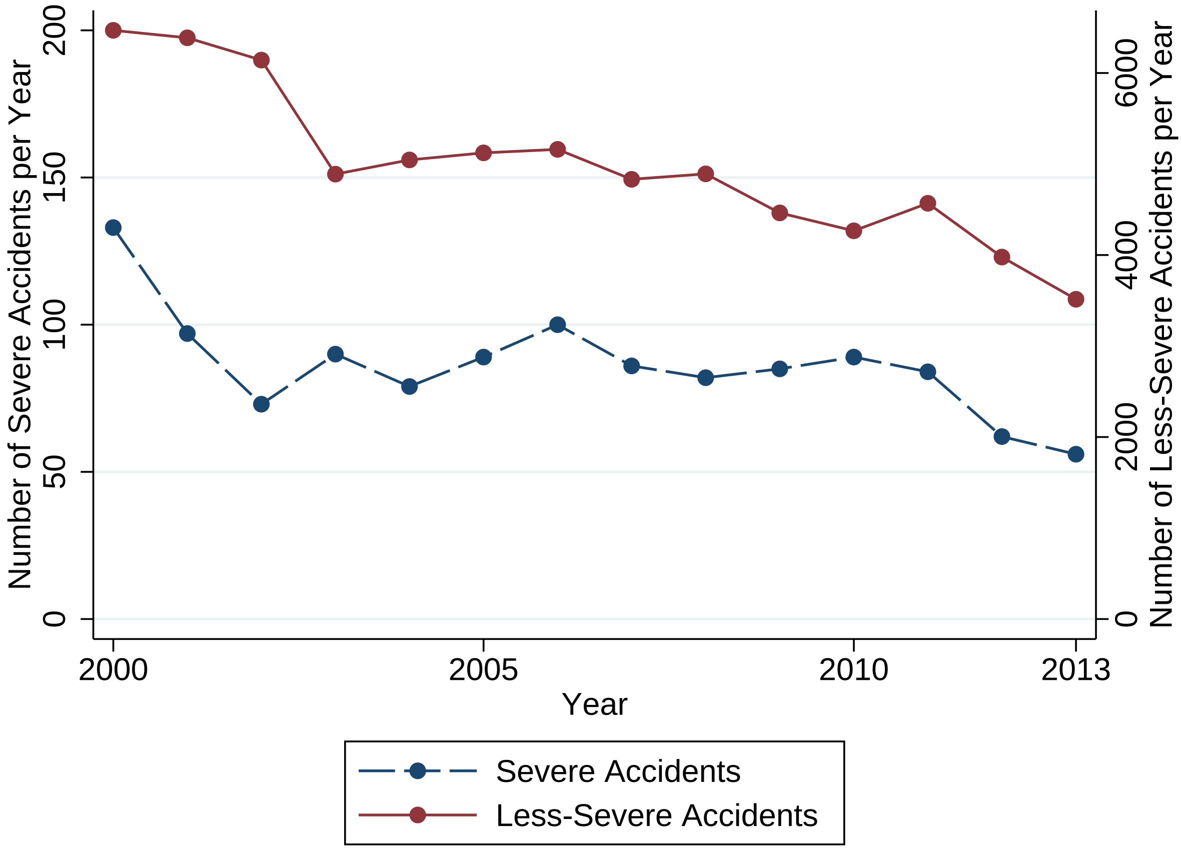


Table 3.2: Summary Statistics at Mine-Quarter Level

| Variable | Mean | Std. Dev. | Mean Within Mine Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Coal production (thousands of tons) | 271 | 1,221 | 86.3 | 0 | 31,354 |
| Productivity (tons per worker hour) | 3.9 | 5.2 | 1.5 | 0 | 87 |
| Hours worked (thousands) | 40 | 64.7 | 13.2 | 2 | 882 |
| Employees | 69 | 113 | 21.3 | 2 | 1,661 |
| Fatalities per million hours | 0.2 | 5.0 | 1.0 | 0 | 467 |
| Severe accidents per million hours | 0.5 | 7.1 | 2.3 | 0 | 467 |
| Less-severe accidents per million hours | 28 | 61.3 | 44.4 | 0 | 2,238 |
| MSHA inspections | 3.5 | 5.7 | 2.0 | 0 | 58 |
| MSHA inspection hours | 166 | 272 | 87.1 | 0 | 8,526 |
| MSHA penalties (thousands of $) | 14 | 56.7 | 19.4 | 0 | 1,982 |
| MSHA violations | 17 | 30.6 | 11.7 | 0 | 470 |

Note: summary statistics are for estimation sample for specifications that have disasters as the main regressor. Sample period is Q1:2000 through Q3:2012. N=51,477. See text for details of sample construction and variable definitions.

person hours, which is the equivalent of four full-time employees over the quarter. Last, as noted above, we exclude from our sample the five mines which experienced a disaster as these mines may be physically damaged by the disaster.

Our data extend from Q1:2000 through Q3:2014. Our regressors include lags and leads of disasters and fatalities up to two years. This limits our sample that uses fatalities to Q1:2002 through Q3:2012. However, because mine disasters are large, public event, we know when they occurred. The most recent mine disaster prior to our sample was in 1992, more than two years preceding our earliest data. Thus, for our regressions which include mine disasters as the main regressor, we can keep observations going back to Q1:2000.

Table 3.2 provides summary statistics on our estimation sample at the mine-quarter level. We include here all observations from Q1:2000 through Q3:2012, which corresponds to the sample with disasters as the main regressor; the sample with fatalities as the main regressor starts in Q1:2002. The mean hours worked is about 40,000, reflecting a mean mine size with a full-time equivalent of 80 workers. The largest mine is about 20 times this size. Mean coal production is about 270,000 tons per quarter. Mean productivity is 3.9 tons.[9]

Despite the high absolute numbers noted in Table 3.1, severe accidents are rare, with a rate of 0.5 per million hours worked. One million hours worked corresponds to 2,000 people working full-time over a quarter. Less-severe accidents occur an

[9] MSHA (and our paper) use the short ton, which is equal to 2,000 pounds.

average of 27.8 times per million hours worked, suggesting that a larger mine with 1,000 workers would expect to have 13.9 such less-severe accidents each quarter. MSHA inspects mines 3.5 times per quarter on average spending an average of 166.2 hours on their inspections. They also assess penalties of $14,100 on average per mine-quarter, finding 16.8 violations on average.

For each reported statistic, the standard deviation is larger than the mean. This indicates that there is substantial variation in all the variables in our data. This variation reflects the diversity of mining operations in the United States, which vary in size from small mines employing a handful of workers to massive sites employing over a thousand workers. These operations likely have different rates of compliance with safety regulations.

Last, Table 3.2 provides information about within-mine variation by showing the mean of the within-mine standard deviations. This statistic is important because our identification of the response to exogenous shocks relies on within-mine variation given that we include mine fixed effects. As the table shows, there is also substantial within-mine variation for all of our variables of interest in our sample. For instance, the within-mine standard deviations in productivity has a mean of 1.5 tons per worker hour, compared to the unconditional standard deviation of 5.2. MSHA enforcement activity variables all have relatively large within-mine standard deviation. For instance, the within-mine standard deviation in the number of MSHA inspections per quarter has a mean of 2.0 compared to the unconditional standard deviation of 5.7.

In addition to MSHA data, we also obtained U.S. Census data from the Integrated Public Use Microdata Series (IPUMS, Ruggles et al., 2010). We use IPUMS data from the American Community Survey, which began in 2005. For our purposes, the IPUMS data list worker occupations for a sample of workers employed in the coal mining industry. These data allow us to understand variation in the effects of a disaster in a state by occupation. Unlike the MSHA data, which are at the mine-quarter level, the American Community Survey data are effectively only usable at the state-year level.

We extracted data from IPUMS for all employees who report coal mining as their occupation, using the 1990 industry code 41. We use the “occ1990” field as our measure of occupation. We split coal-mining workers into three occupations based on this field: coal miners, managers (defined as individuals whose occupation includes

Table 3.3: Summary Statistics for Coal Mining Occupations at State-Year Level

| Worker Occupation | Observations | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Miners | 220 | 1,111 | 1,693 | 0 | 8,913 |
| Managers | 220 | 603 | 906 | 0 | 4,263 |
| Other Workers | 220 | 2,932 | 4,463 | 0 | 22,615 |
| All | 220 | 4,647 | 6,931 | 22 | 33,561 |

Note: data are obtained from IPUMS, extend from 2005 through 2013.

the word "manager" or "supervisor"), and other workers.

The IPUMS data include survey weights and we collapse the data to the state and year level using these weights. The data extend from 2005 through 2013, a shorter time period than our overall study period, thus limiting the power of this analysis.

Table 3.3 provides summary statistics on the IPUMS data. Our IPUMS data include 220 state-year observations, all of which have some workers in the coal mining industry. Miners comprise about 24% of the employees in this sector, managers 13%, and other occupations 63%. The state with the most

## 3.5 Results

### 3.5.1 Effect of Mine Fatalities

We start by examining the impact of mine fatalities on the mine itself. We first report on the "first stage" of the impact of MSHA enforcement activity. Here, we report results from a series of regressions that regress different MSHA enforcement actions on the presence of a current and lagged fatality at the mine. All regressions are at the mine-quarter level. We believe that the presence of a fatality represents a quasi-experimental source of variation. In order to partially test this hypothesis, all specifications also include *future* fatality measures as a falsification test. Because larger mines provide more information, we weight all regressions by the mean number of hours worked in that mine. We also include mine and quarter fixed effects and cluster standard errors at the mine level.

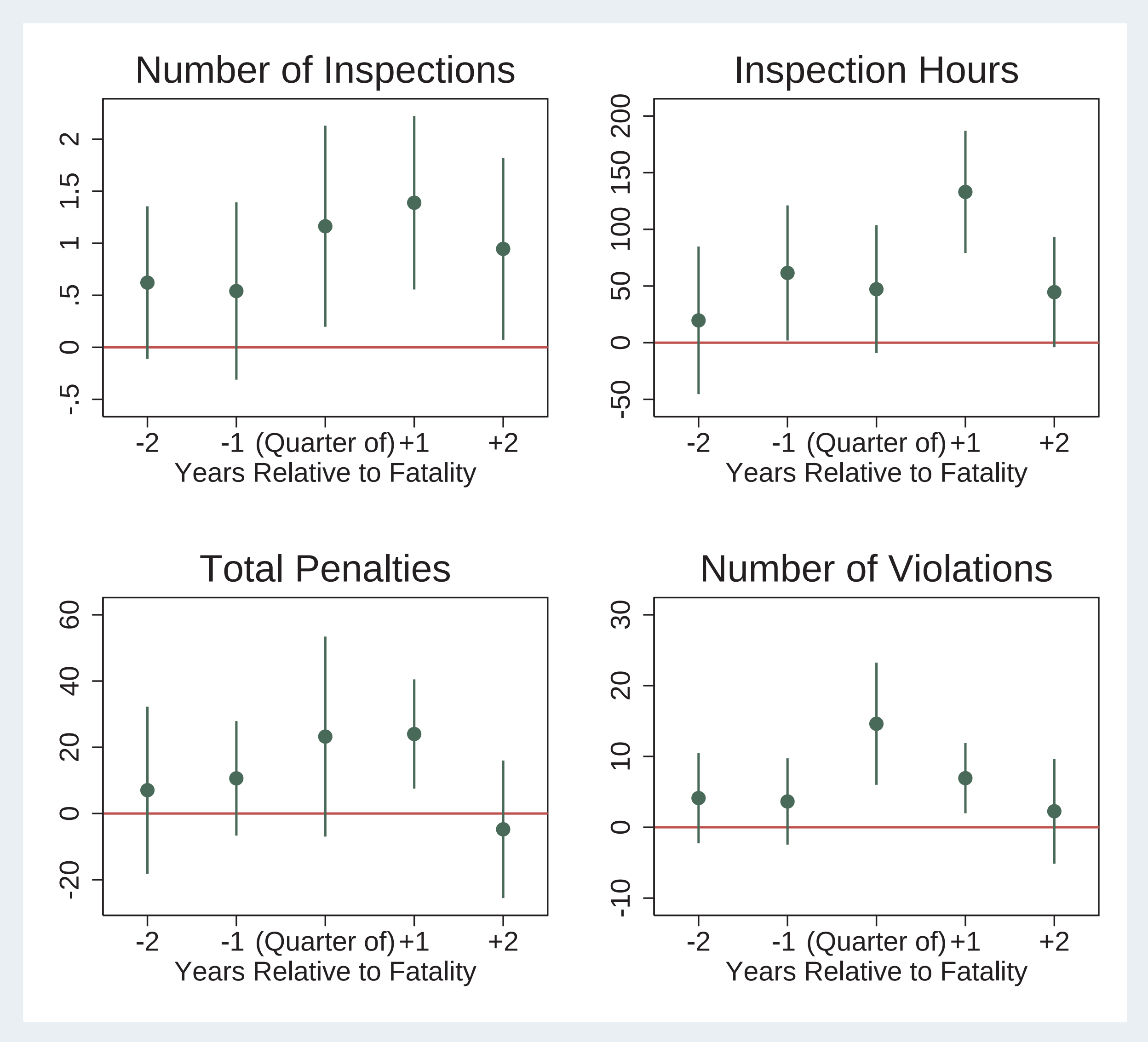
We report the same set of results in two ways. First, Figure 3.4 shows the

Table 3.4: Regressions of Fatalities on MSHA Activity

| | (1) Number of Inspections | (2) Inspection Hours | (3) Penalties (Thousands $) | (4) Number of Violations |
|---|---|---|---|---|
| *Effect by Year* | | | | |
| Two Years Prior to Fatality | 0.62* | 19.7 | 7.04 | 4.13 |
| | (0.37) | (33.2) | (12.9) | (3.26) |
| One Year Prior | 0.54 | 61.5** | 10.6 | 3.64 |
| | (0.43) | (30.4) | (8.81) | (3.11) |
| Quarter of Fatality | 1.16** | 47.2 | 23.2 | 14.6*** |
| | (0.49) | (28.8) | (15.4) | (4.40) |
| One Year After | 1.39*** | 133.0*** | 24.0*** | 6.93*** |
| | (0.43) | (27.5) | (8.41) | (2.53) |
| Two Years After | 0.95** | 44.6* | -4.77 | 2.27 |
| | (0.45) | (24.8) | (10.6) | (3.78) |
| State×Hours Worked | Yes | Yes | Yes | Yes |
| State×Employees | Yes | Yes | Yes | Yes |
| N | 43,377 | 43,377 | 43,377 | 43,377 |
| $R^2$ Adj. | 0.16 | 0.31 | 0.23 | 0.28 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. Dependent variable is indicated in the column title. All regressions include mine fixed effects and quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. Standard errors, reported in parentheses, are clustered at the mine level.

Figure 3.4: Regressions of Fatalities on MSHA Activity




Each figure reports selected regressors from one regression. Dependent variables are indicated in the title. All regressions include mine fixed effects, quarterly fixed effects, and state dummies times hours worked and state dummies times employees (to control for economies of scale) as regressors and are weighted by mean hours worked at the mine. Standard errors are clustered at the mine level. Each line shows a 95% confidence interval.

coefficients on current, future, and lagged fatalities graphically, along with standard errors. Second, Table 3.4 shows the coefficients in table form. We present all results in this same manner.

We find that MSHA inspections, inspection hours, and the number of violations (though not penalties) increase significantly in the four quarters after a fatality at a mine. The magnitudes of the increases are large. For instance, the number of inspections increases by 1.5 in quarters 1-4 after the fatality, representing a 18% increase relative to the baseline at the affected mines at those time periods, while the number of cited violations increases by 6.2, representing a 16% increase.

Interestingly, in the period two years (5-8 quarters) after the fatality, the effects are somewhat different. While the number of inspections and inspection hours are still significantly higher than the baseline, the total penalties and number of violations both fall back to having much smaller coefficients that are statistically no different than the baseline. We believe that this may be caused by firms reacting to the fatality and the increased MSHA enforcement activity by increasing their safety input.

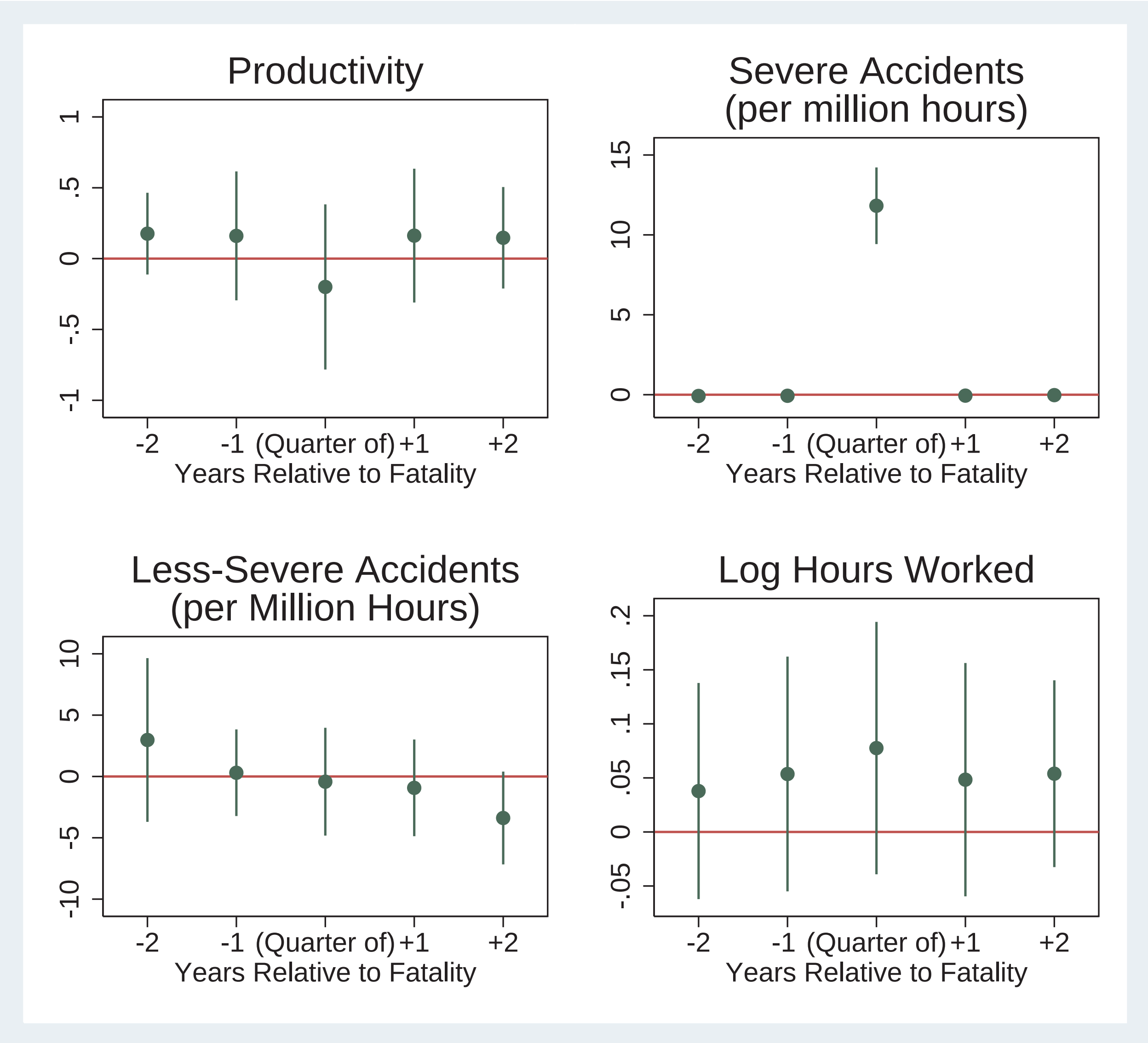
In contrast to the coefficients in the quarter of the fatality and following the fatality, the future presence of a fatality does not significantly affect MSHA enforcement activity, except for one variable in one year (inspection hours in the year before the fatality, with P=0.043).[10] Thus, the data do not seem to indicate the presence of a reverse causality story where common factors are driving both increased MSHA enforcement activity and more fatalities.

Having established that enforcement activity is a potential causal pathway by which fatalities might affect the relative price of safety and mineral output, we next seek to understand the impact of fatalities on productivity and measures of safety. Figure 3.5 and Table 3.5 present results that are analogous to Figure 3.4 and Table 3.4 respectively, regressing productivity and safety measures on current, future, and lagged fatalities. We include the same regressors as in the earlier specifications, except that our specifications with log hours as the dependent variable exclude indicators for hours worked and employees.

Here, we find that the rate of less-severe accidents per hour drop in quarters 5-8 (two years after) the fatality, with marginal significance (P=0.081). The drop is about 7% of the baseline value. The coefficient on severe accidents, though not significant, is also negative. None of the other coefficients on accidents are significant, except for the coefficient on accidents in the quarter of fatality, which is significantly positive and large, reflecting the fact that the fatality itself is a severe accident. We find no impact on productivity or log hours. We find no support for the reverse causality story.

[10] Even in this case, the two dummies, on one and two years prior to the fatality, are not jointly significant (P=0.125).

Figure 3.5: Regressions of Fatalities on Productivity and Safety




Each figure reports selected regressors from one regression. Dependent variables are indicated in the title. All regressions include mine fixed effects, quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. All regressions but log hours include state dummies times hours worked and state dummies times employees a regressors. Standard errors are clustered at the mine level. Each line shows a 95% confidence interval.

Together, the results appear to support our hypothesis that mine fatalities are relatively rare events that shift the relative price of safety and mineral output for a mine through increased enforcement activity. The increased price of safety appears to cause mines to decrease their less-severe accidents by 7%. This does not appear to have any adverse impact of productivity, and indeed the point estimates on productivity and hours worked are positive.

Table 3.5: Regressions of Fatalities on Productivity and Safety

| | (1) Less-Severe Accidents (per million hours) | (2) Severe Accidents (per million hours) | (3) Productivity | (4) Log Worker Hours |
|---|---|---|---|---|
| *Effect by Year* | | | | |
| Two Years Prior to Fatality | 2.97 | -0.077 | 0.18 | 0.038 |
| | (3.40) | (0.089) | (0.15) | (0.051) |
| One Year Prior | 0.30 | -0.066 | 0.16 | 0.054 |
| | (1.80) | (0.14) | (0.23) | (0.055) |
| Quarter of Fatality | -0.43 | 11.8*** | -0.20 | 0.078 |
| | (2.24) | (1.22) | (0.30) | (0.060) |
| One Year After | -0.93 | -0.058 | 0.16 | 0.048 |
| | (2.01) | (0.14) | (0.24) | (0.055) |
| Two Years After | -3.39* | -0.026 | 0.15 | 0.054 |
| | (1.93) | (0.11) | (0.18) | (0.044) |
| State×Hours Worked | Yes | Yes | Yes | No |
| State×Employees | Yes | Yes | Yes | No |
| N | 43,377 | 43,377 | 43,377 | 43,377 |
| $R^2$ Adj. | 0.05 | 0.09 | 0.27 | 0.03 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. Dependent variable is indicated in the column title. All regressions include mine fixed effects and quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. Standard errors, reported in parentheses, are clustered at the mine level.

### 3.5.2 Effect of Mine Disasters

We now present the effect of mine disasters, defined as five deaths or more at a mine. Recall that our sample includes five disasters (Figure 3.2). Because a disaster may permanently shut down or alter a mine, we are not as interested in the effect of a disaster on the mine itself. Rather, we are interested in understanding whether the disaster has an effect on mines in the nearby area although we also examine measures based on a fixed distance. As noted in the introduction, this may occur through media or regulatory attention or increased worker concern, among other factors. Our base results consider the impact of mines within the same state, where state is proxying for the nearby area. Of course, some of the impact of a mine disaster may be national, as evidenced by President Obama's speech noted in the introduction. Our estimation will capture national effects of the disaster through the time dummies and hence our estimate of the impact of the disaster will only capture the part of the effect that is local.

Figure 3.6 and Table 3.6 present evidence for the effects of disasters on MSHA enforcement activity. For each of the four enforcement variables, we find no pattern of significant variation in MSHA enforcement activity in mines following a mine disaster for mines in the same state as the disaster. Thus, in the case of mine disasters, to the extent that disasters change the relative price of safety and mineral output, the causal pathway does not appear to be MSHA enforcement activity.

We next turn to the effect of disasters on productivity and measures of safety. Figure 3.7 and Table 3.7 present these results. We find that the rates of fatalities and less-severe accidents per hour both drop steeply and significantly two years after a disaster in the state. The results are large with a coefficient of $-0.22$ on fatalities and $-8.1$ on less-severe accidents. These imply a drop of 68% in fatalities per hour worked and 23% in less-severe accidents per hour relative to the baseline at the affected mines at those time periods. As a caveat, the rates of these accidents were both negative preceding the disaster. However, the coefficients were generally much smaller and not statistically significant prior to the disaster. Even using the future rate as the baseline would imply a big drop in both severe and less-severe accidents from a disaster in the state. The impact of a disaster on severe accidents is not statistically significant but is similar in magnitude to the impact on fatalities.

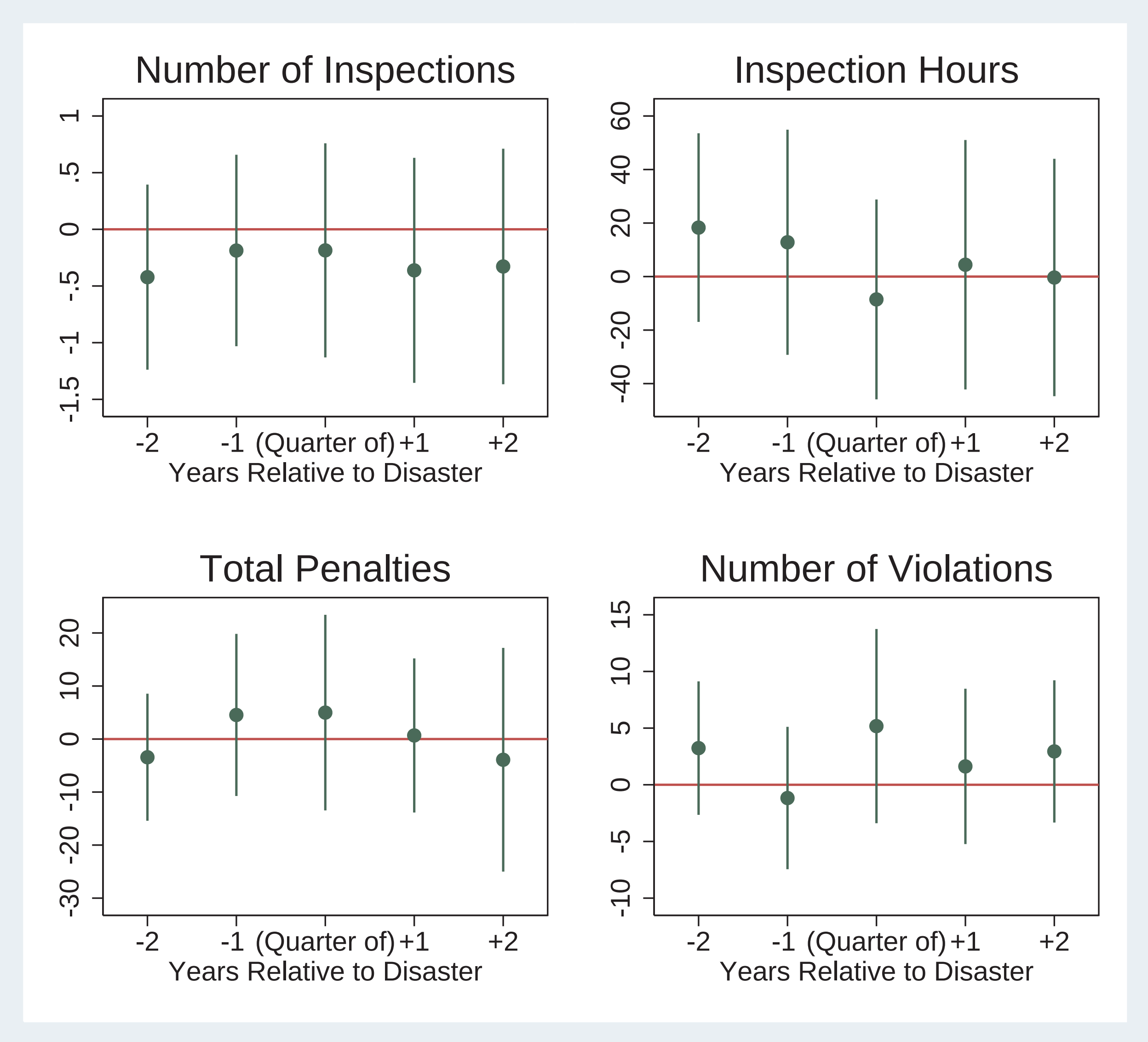
Overall, we believe that there is strong evidence that a disaster in a state causes

Table 3.6: Regressions of Disasters on MSHA Activity

| | (1) Number of Inspections | (2) Inspection Hours | (3) Penalties (Thousands $) | (4) Number of Violations |
|---|---|---|---|---|
| *Effect by Year* | | | | |
| Two Years Prior to Disaster | -0.42 | 18.3 | -3.43 | 3.24 |
| | (0.42) | (18.0) | (6.11) | (3.00) |
| One Year Prior | -0.19 | 12.8 | 4.54 | -1.17 |
| | (0.43) | (21.5) | (7.80) | (3.20) |
| Quarter of Disaster | -0.19 | -8.54 | 4.99 | 5.18 |
| | (0.48) | (19.0) | (9.41) | (4.37) |
| One Year After | -0.36 | 4.42 | 0.68 | 1.62 |
| | (0.51) | (23.8) | (7.41) | (3.50) |
| Two Years After | -0.33 | -0.36 | -3.91 | 2.94 |
| | (0.53) | (22.6) | (10.8) | (3.20) |
| State×Hours Worked | Yes | Yes | Yes | Yes |
| State×Employees | Yes | Yes | Yes | Yes |
| N | 51,477 | 51,477 | 51,477 | 51,477 |
| $R^2$ Adj. | 0.18 | 0.34 | 0.25 | 0.31 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. Dependent variable is indicated in the column title. All regressions include mine fixed effects and quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. Standard errors, reported in parentheses, are clustered at the mine level.

Figure 3.6: Regressions of Disasters on MSHA Activity

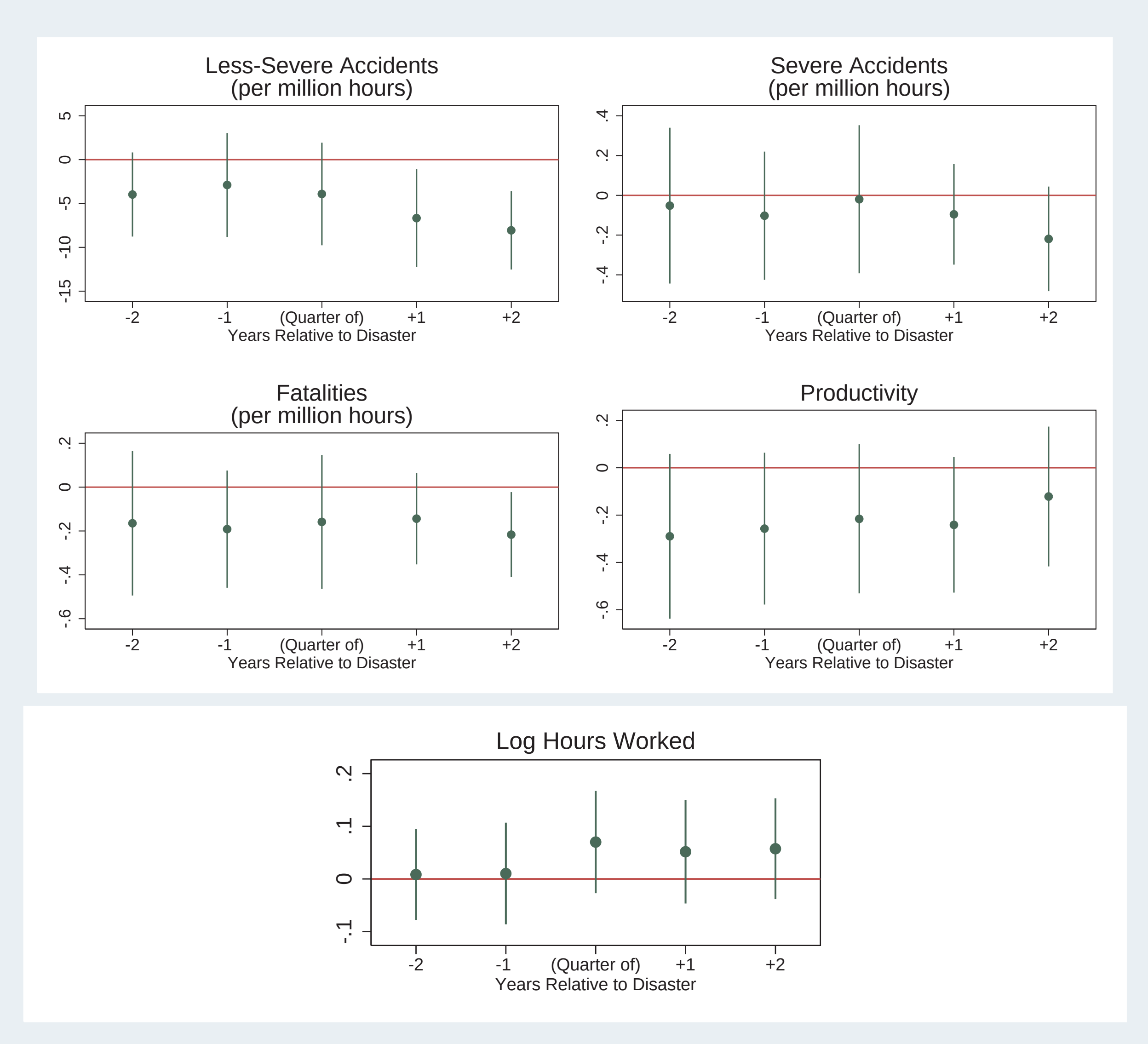



Each figure reports selected regressors from one regression. Dependent variables are indicated in the title. All regressions include mine fixed effects, quarterly fixed effects, and state dummies times hours worked and state dummies times employees as regressors and are weighted by mean hours worked at the mine. Standard errors are clustered at the mine level. Each line shows a 95% confidence interval.

mines to increase safety inputs. Note also that the magnitude of the safety effects are much bigger than we find after a single fatality.

We find weaker evidence that mineral productivity drops following a disaster in the state: the coefficient in the year after the disaster is negative at $-0.24$ and marginally significant with $P = 0.098$. Moreover, even though future disasters do not

Figure 3.7: Regressions of Disasters on Productivity and Safety




Each figure reports selected regressors from one regression. Dependent variables are indicated in the title. All regressions include mine fixed effects, quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. All regressions but log hours include state dummies times hours worked and state dummies times employees a regressors. Standard errors are clustered at the mine level. Each line shows a 95% confidence interval.

significantly predict drops in productivity, the magnitude of the future coefficients is very similar (and actually slightly larger than) the $-0.24$ coefficient here. Although the model predicted that we should see a decrease in hours worked, we do not see any significant effect on hours worked following the disaster. Thus, there is at best very weak evidence that our rational model of a tradeoff between productivity and safety is an accurate description of safety behavior for coal mines.

Table 3.7: Regressions of Disasters on Productivity and Safety

| | (1) Less-Severe Accidents (per million hours) | (2) Severe Accidents (per million hours) | (3) Fatalities (per million hours) | (4) Productivity | (5) Log Worker Hours |
|---|---|---|---|---|---|
| *Effect by Year* | | | | | |
| Two Years Prior to Disaster | -3.98 | -0.052 | -0.16 | -0.29 | 0.0084 |
| | (2.45) | (0.20) | (0.17) | (0.18) | (0.044) |
| One Year Prior | -2.89 | -0.10 | -0.19 | -0.26 | 0.010 |
| | (3.02) | (0.16) | (0.14) | (0.16) | (0.049) |
| Quarter of Disaster | -3.91 | -0.020 | -0.16 | -0.22 | 0.070 |
| | (2.98) | (0.19) | (0.16) | (0.16) | (0.049) |
| One Year After | -6.67** | -0.095 | -0.14 | -0.24* | 0.052 |
| | (2.84) | (0.13) | (0.11) | (0.15) | (0.050) |
| Two Years After | -8.06*** | -0.22 | -0.22** | -0.12 | 0.057 |
| | (2.28) | (0.13) | (0.099) | (0.15) | (0.049) |
| State×Hours Worked | Yes | Yes | Yes | Yes | No |
| State×Employees | Yes | Yes | Yes | Yes | No |
| N | 51,477 | 51,477 | 51,477 | 51,477 | 51,477 |
| $R^2$ Adj. | 0.04 | 0.00 | 0.00 | 0.31 | 0.03 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. Dependent variable is indicated in the column title. All regressions include mine fixed effects and quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. Standard errors, reported in parentheses, are clustered at the mine level.

Our results here focus on the effect of mine disasters on other mines in the same state. However, we have also examined other measures of closeness, such as mines located within 200 KM and 400 KM. These results, which we present in Appendix A, are very similar to our base results in Figure 3.7 and Table 3.7.

We also estimated specifications of the impact of disasters on productivity and safety with continuous measures of distance. Here, we present results that allow for a base effect of being within 200 KM of a disaster and a term that interacts distance to the disaster with being within the 200 KM threshold. Table 3.8, which presents these results, shows base effects that confirm our results in Table 3.7: fatalities, severe accidents, and less-severe accidents all decrease after the disaster, but productivity and hours worked do not change significantly.

However, the distance effects also show some interesting differences. For both fatalities per hour worked and severe accidents per hour worked, the impact of the disaster is not significantly different the further one is from the disaster. But, the impact of the disaster on reducing less-severe accidents decays the further one is from the mine with the disaster. The decay effect is large, with a mine collocated at the disaster experiencing $-9.15$ fewer less-severe accidents two years after the disaster, but the effect being only $-1.49$ at 200 KM from the disaster. One possible explanation is that while mine owners face strong incentives to reduce fatalities and severe accidents in the wake of a disaster, they face lower incentives to reduce less-severe accidents. Instead, reductions in these relatively minor accidents are produced by face to face communication and word of mouth among non-executive employees across mines. Another explanation is that mines closer to the disaster are more likely to have in common the same types of equipment and procedures and are better able to share improved methods to reduce less-severe accidents.

Finally, we examine how disasters in a state change the composition of employees, using the American Community Survey data from IPUMS and regressions at the state-year level. Since our data here pertain only to 2005-13, they contain only four disasters. Moreover, we omit the falsification indicator for being two years before a disaster (but include the indicator for being one year before a disaster), as otherwise, we would be reduced to having two disasters in our sample.[11] We find that in the year of a disaster and the year following a disaster, there are fewer miners in the state of the disaster, but that this number rebounds to roughly the baseline level two years after the disaster. We also find that the number of other workers is

[11]Also, unlike our base regressions, these regressions do not exclude the mine with the disaster.

Table 3.8: Regressions of Disasters on Productivity and Safety, With Distance Interactions

empty

| | (1) Less-Severe Accidents (per million hours) | (2) Severe Accidents (per million hours) | (3) Fatalities (per million hours) | (4) Productivity | (5) Log Worker Hours |
|---|---|---|---|---|---|
| Two Years Prior to Disaster | | | | | |
| Base | -2.85 | -0.33 | -0.39* | -0.34 | -0.020 |
| | (3.63) | (0.25) | (0.23) | (0.22) | (0.042) |
| Base × | 1.30 | 0.14 | 0.15** | 0.050 | -0.0082 |
| Distance (per 100km) | (2.59) | (0.12) | (0.073) | (0.091) | (0.033) |
| One Year Prior | | | | | |
| Base | -5.84 | 0.0092 | -0.33 | -0.16 | -0.076 |
| | (3.64) | (0.29) | (0.22) | (0.21) | (0.055) |
| Base × | 2.42 | -0.31* | 0.064 | -0.029 | 0.019 |
| Distance (per 100km) | (2.20) | (0.19) | (0.074) | (0.093) | (0.040) |
| Quarter of Disaster | | | | | |
| Base | -3.46 | 0.077 | -0.30 | -0.11 | -0.071 |
| | (3.79) | (0.31) | (0.26) | (0.21) | (0.066) |
| Base × | 0.21 | -0.20 | 0.16 | -0.014 | 0.037 |
| Distance (per 100km) | (2.50) | (0.23) | (0.17) | (0.11) | (0.048) |
| One Year After | | | | | |
| Base | -7.08** | -0.34* | -0.35** | -0.060 | -0.055 |
| | (3.24) | (0.19) | (0.18) | (0.20) | (0.068) |
| Base × | 3.12* | -0.031 | 0.048 | -0.085 | 0.0088 |
| Distance (per 100km) | (1.78) | (0.081) | (0.045) | (0.092) | (0.046) |
| Two Years After | | | | | |
| Base | -9.15*** | -0.37* | -0.33** | 0.060 | -0.011 |
| | (2.50) | (0.20) | (0.16) | (0.19) | (0.060) |
| Base × | 3.83** | 0.019 | 0.086 | -0.032 | -0.027 |
| Distance (per 100km) | (1.80) | (0.11) | (0.063) | (0.076) | (0.037) |
| State× Hours Worked | Yes | Yes | Yes | Yes | No |
| State×Employees | Yes | Yes | Yes | Yes | No |
| N | 51,477 | 51,477 | 51,477 | 51,477 | 51,477 |
| $R^2$ Adj. | 0.04 | 0.00 | 0.00 | 0.31 | 0.03 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. The units of the Base × Distance regressor is in 100km. The dependent variable is indicated in the column title. All regressions include mine fixed effects and quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. Standard errors, reported in parentheses, are clustered at the mine level.

Table 3.9: Regressions of Disasters on the Number of Employees by Occupation

| | (1) All Workers | (2) Managers | (3) Miners | (4) Other Workers |
|---|---|---|---|---|
| One Year Prior to Disaster | -1315.2* | -102.9 | -911.9 | -300.4 |
| | (742.8) | (387.2) | (626.3) | (949.2) |
| Year of Disaster | 1642.3*** | 455.5** | -663.0 | 1849.8** |
| | (538.1) | (173.6) | (512.4) | (747.5) |
| One Year After | -763.5 | -58.8 | -954.2*** | 249.4 |
| | (648.0) | (240.7) | (173.3) | (908.5) |
| Two Years After | 1679.1*** | 252.0*** | 533.4 | 893.7 |
| | (386.3) | (61.0) | (804.3) | (1035.1) |
| N | 220 | 220 | 220 | 220 |
| $R^2$ Adj. | 0.51 | 0.26 | 0.33 | 0.27 |

empty

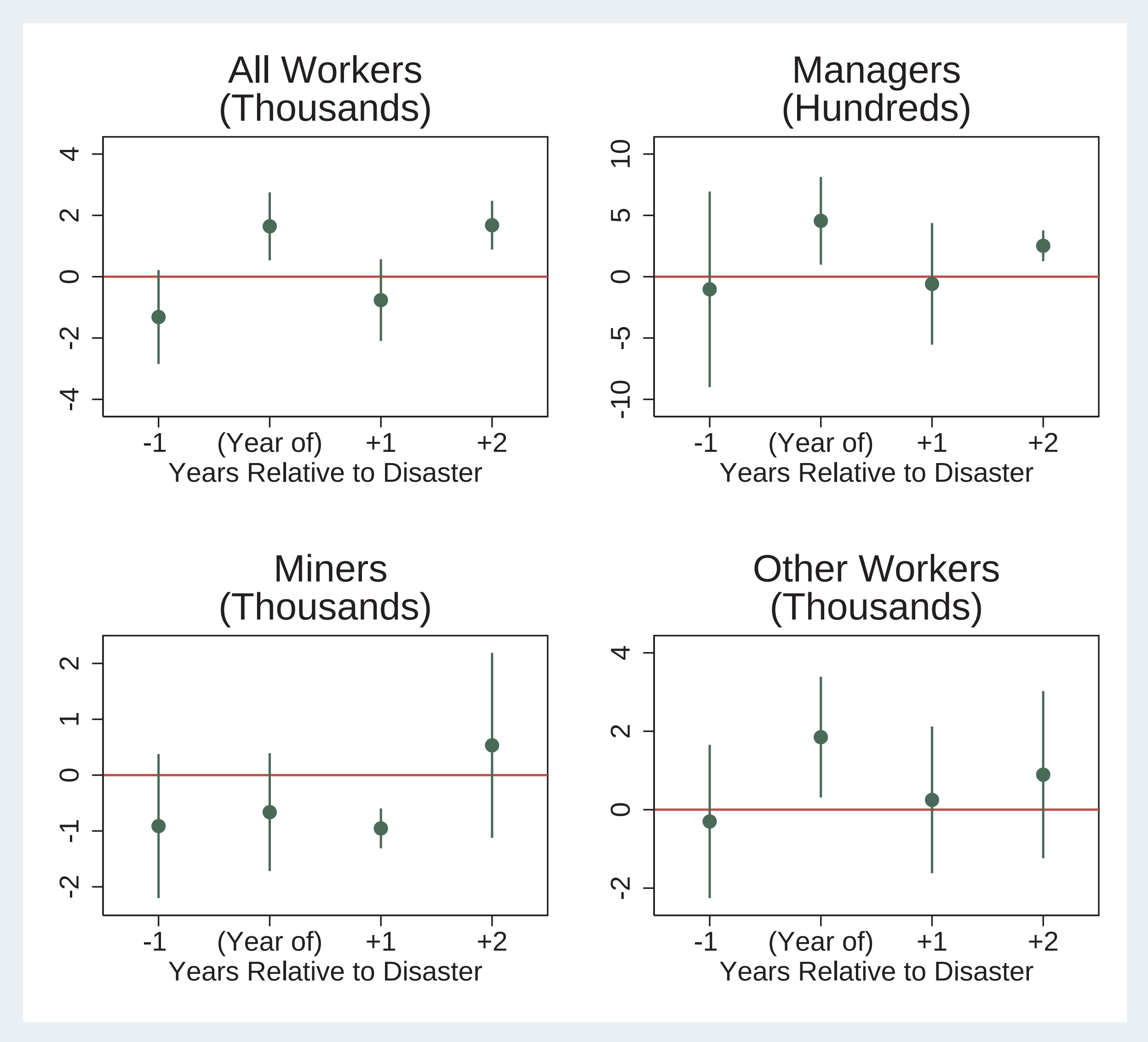
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. The dependent variable is indicated in the column title. All regressions include state and year fixed effects as regressors and are weighted by the mean number of workers in the state. Standard errors, reported in parentheses, are clustered at the state level.

significantly higher in the year and state of the disaster, but that this figure also is not significantly different than the baseline in the years following a disaster. In contrast, the number of managers increases by 252 two years after a disaster, which represents an 11% increase relative to the baseline at the affected states at those time periods. Unlike in the base regressions, the overall number of workers increases two years after a disaster in the state. The extra 1,679 workers represents an 8% increase relative to the baseline. Thus, it is possible that firms employ more managers and supervisors two years after a disaster in their state with the goal of reducing accident risk. The hiring of additional personnel who do not produce coal would cause productivity decreases.

### 3.5.3 Dollar Magnitudes of Effects

We now turn to evaluating the dollar magnitudes of our effects. We first examine the dollar cost savings that mines may incur from the lowered accident rates following the disaster. To the extent that the increased safety input lowers the risk of fatality, there is a large literature on the value of a statistical life (VSL) that seeks to estimate the cost of fatalities. A number of papers in this literature use labor market data to estimate the wage premiums that workers earn from dangerous occupations

Figure 3.8: Regressions of Disasters on the Number of Employees, by Category




Each figure reports selected regressors from one regression. Dependent variables are indicated in the title. All regressions include state and year fixed effects and are weighted by the mean number of workers in the state. Standard errors are clustered at the state level. Each line shows a 95% confidence interval.

such as mining, and divide the wage premiums by the probability of death. The fatality cost measured by these papers are costs borne by firms, who pay higher wages for dangerous occupations. While this context of estimating the VSL is the most similar to our paper, the VSL literature does not account for the direct payouts that firms make to workers who are injured or to MSHA for cited safety violations. Moreover, the magnitude of these payouts may be large. The National Safety

Council estimates a workplace fatality to cost $1.4 million (National Safety Council, 2014), while BP reportedly paid over $8 million for each fatality that resulted from the Deepwater Horizon oil-rig explosion (Dionne, 2011).

An influential review article, Viscusi and Aldy (2003), finds that estimates of the VSL that use U.S. labor market data are between $4 and $9 million. Using the midpoint value of $6.5 million (Viscusi and Aldy, 2003), we find that the reduction in risk of fatalities two years after the disaster is worth $1.41 per hour worked.

We next evaluate the cost savings from lowering less-severe accidents. The National Safety Council estimates that the average cost of an accident which results in work absence is $30,000 (National Safety Council, 2014). Injuries resulting in work absence are our most common type of less-severe accidents (Table 3.1). This figure, together with our estimates above, imply a cost savings of $0.24 per hour worked from the reduction in less-severe accidents.

Neither of these figures accounts for the reduced costs of regulatory compliance, MSHA violations, or potential reductions in accidents with severe injuries other than fatalities. Indeed, Viscusi and Aldy (2003) state that permanent disabilities should be valued similarly to fatalities. Overall, we believe that $1-$2 per hour worked is a reasonable estimate for the total dollar cost savings to the firm from the decreased accidents following a mine disaster.

Although we do not believe that there is strong evidence of a productivity loss following a disaster, it is worth noting that the $0.024$ drop in tons of coal per hour worked noted above, if real, would imply a 7% drop in productivity, implying the need to add an extra 8% work hours to mine the same coal. At $25/hour,[12] this represents an extra $2 in wages per current hour worked. Moreover, this may be a serious underestimate of the loss from the productivity drop. Coal sells for about $50/ton[13] and workers produce just under four tons of coal per hour on average (Table 3.2). Thus, to the extent that firms simply produce less following a disaster rather than being able to hire more safety-related workers without capital costs, the reduction in coal productivity would cost the firm about $14 per hour worked.

[12] $25 is close to the mean wage for the majority of workers that are being exposed to accident risk. BLS reports the three largest occupations and their wages to be "Construction and Extraction" with mean wage of $25.14, "Extraction Workers" with a mean wage of $25.10 and "Transportation and Material Moving Occupations" with a mean wage of $23.61. Virtually all other non-supervisor or non-management occupations have wages with a similar value. See Table NAICS 212100 in U.S. Bureau of Labor Statistics (2014).

[13] The average sales price of coal in 2013 was $37.24 per ton, but sales prices from underground mines were higher at $60.98, and in each state affected by a disaster, the average price of coal is generally higher than this average (Alabama $88, Kentucky $59, West Virginia $75, Utah $35. See Table 28 in Energy Information Administration (2015).)

Although we do not believe that our productivity numbers are at all conclusive, it is very possible that there are productivity losses following a disaster that cost the firm more than the savings from the increased safety using values of fatalities from the VSL literature. This might be because the VSL literature only accounts for one component of the cost of accidents, which is higher wages, but does not account for other costs that firms pay following disasters including stigma from bad publicity. This, the published VSL figures may substantially underestimate the cost to the firm from accident risk.

## 3.6 Conclusion

Coal mining remains a dangerous occupation where firms and workers may be implicitly making tradeoffs between mineral production and safety. Mine regulation through MSHA is substantial in this sector, implying further that these tradeoffs may be determined by the extent of government regulation. This paper seeks to understand these tradeoffs. We hypothesize that a fatality at a coal mine increases the relative cost of a future accident and that a disaster at a coal mine has a broader impact, increasing the relative cost of a future accident in mines located near a disaster. We use fatalities and disasters as sources of quasi-experimental variation that allow us to trace out the production possibility set between mineral output and safety.

We find that fatalities cause a large increase in MSHA enforcement activity. This increase in enforcement activity appears to spur an increase in safety production, with less-severe accidents dropping 10% in the period two years after a fatality. Moreover, this drop in less-severe accidents is not accompanied by a drop in productivity, implying that firms and workers may be able to improve safety at the margin without incurring a drop in revenues from lower mineral production and that regulatory enforcement may be useful in causing this outcome.

In contrast, we find that disasters in the same state do not cause any change in MSHA enforcement activity. Yet, there is plenty of anecdotal evidence that a mine disaster causes media exposure and public pressure for safety. We find that following a mine disaster, there is a much larger drop in fatalities and less-severe accidents. The effect of less-severe accidents appears in mines local to the mine with the disaster while the effect on fatalities and severe accidents occurs as far as 200 KM away. There is also some marginal evidence that this drop may be accompanied by a drop in productivity. Finally, more managers and supervisors are employed in states with a disaster two years after the disaster.

Overall, our results suggest that MSHA enforcement activity may be helpful in affecting safety at the margin but not as much in reducing severe accidents and deaths. Our site visits to mines reinforce these findings, with safety officers reporting that MSHA inspections focus on minor violations (such as open garbage can lids) rather than on systemic issues. In contrast, public pressure from mine disasters may be more useful in changing the culture of safety to increase safety production. Yet, this increase may also cause drops in mineral productivity.

Finally, it appears that coal mines are operating close to the frontier of the mineral output-safety production possibility set. We do not find evidence, such as in Schmitz (2005), Hausman (2014), or Hendel and Spiegel (2014), that coal mines are operating far from the frontier. It is possible that this is because the coal mining industry is more competitive than industries such as steel milling or iron ore processing, since it includes many firms and a sizable export market.

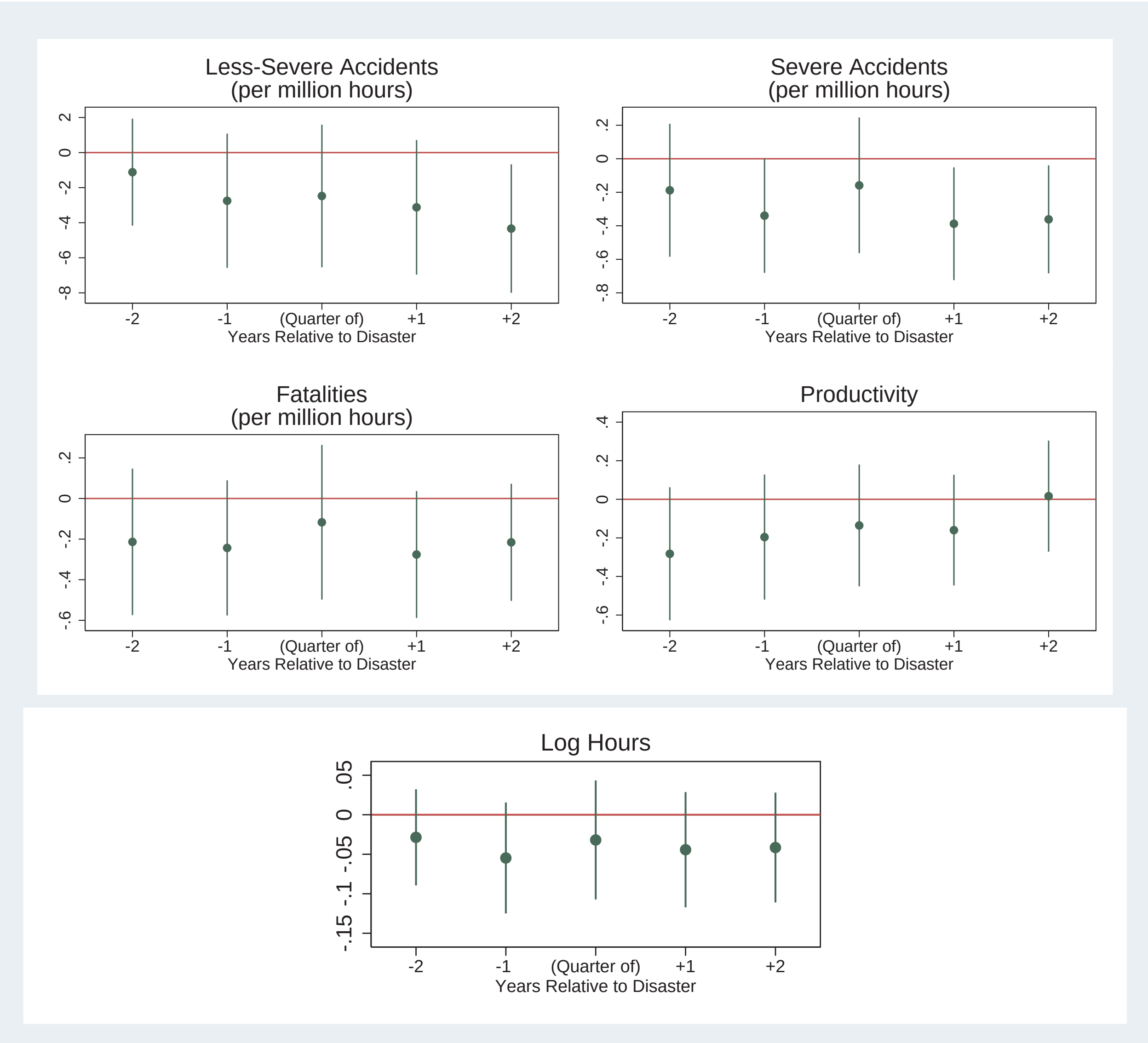
# Appendix 3.A Supplementary Regression Results

Table A1: Regressions of Disasters Within 200 KM on Productivity and Safety

| | (1) Less-Severe Accidents (per million hours) | (2) Severe Accidents (per million hours) | (3) Fatalities (per million hours) | (4) Productivity | (5) Log Worker Hours |
|---|---|---|---|---|---|
| *Effect by Year* | | | | | |
| Two Years Prior to Disaster | -1.12 | -0.19 | -0.21 | -0.28 | -0.029 |
| | (1.56) | (0.20) | (0.18) | (0.18) | (0.031) |
| One Year Prior | -2.75 | -0.34* | -0.24 | -0.20 | -0.055 |
| | (1.95) | (0.17) | (0.17) | (0.17) | (0.036) |
| Quarter of Disaster | -2.48 | -0.16 | -0.12 | -0.14 | -0.032 |
| | (2.07) | (0.21) | (0.19) | (0.16) | (0.038) |
| One Year After | -3.12 | -0.39** | -0.28* | -0.16 | -0.044 |
| | (1.96) | (0.17) | (0.16) | (0.15) | (0.037) |
| Two Years After | -4.34** | -0.36** | -0.22 | 0.016 | -0.041 |
| | (1.87) | (0.16) | (0.15) | (0.15) | (0.035) |
| State× Hours Worked | Yes | Yes | Yes | Yes | No |
| State×Employees | Yes | Yes | Yes | Yes | No |
| N | 51,477 | 51,477 | 51,477 | 51,477 | 51,477 |
| $R^2$ Adj. | 0.04 | 0.00 | 0.00 | 0.31 | 0.03 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. Dependent variable is indicated in the column title. All regressions include mine fixed effects and quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. Standard errors, reported in parentheses, are clustered at the mine level.

Figure A1: Regressions of Disasters Within 200 KM on Productivity and Safety



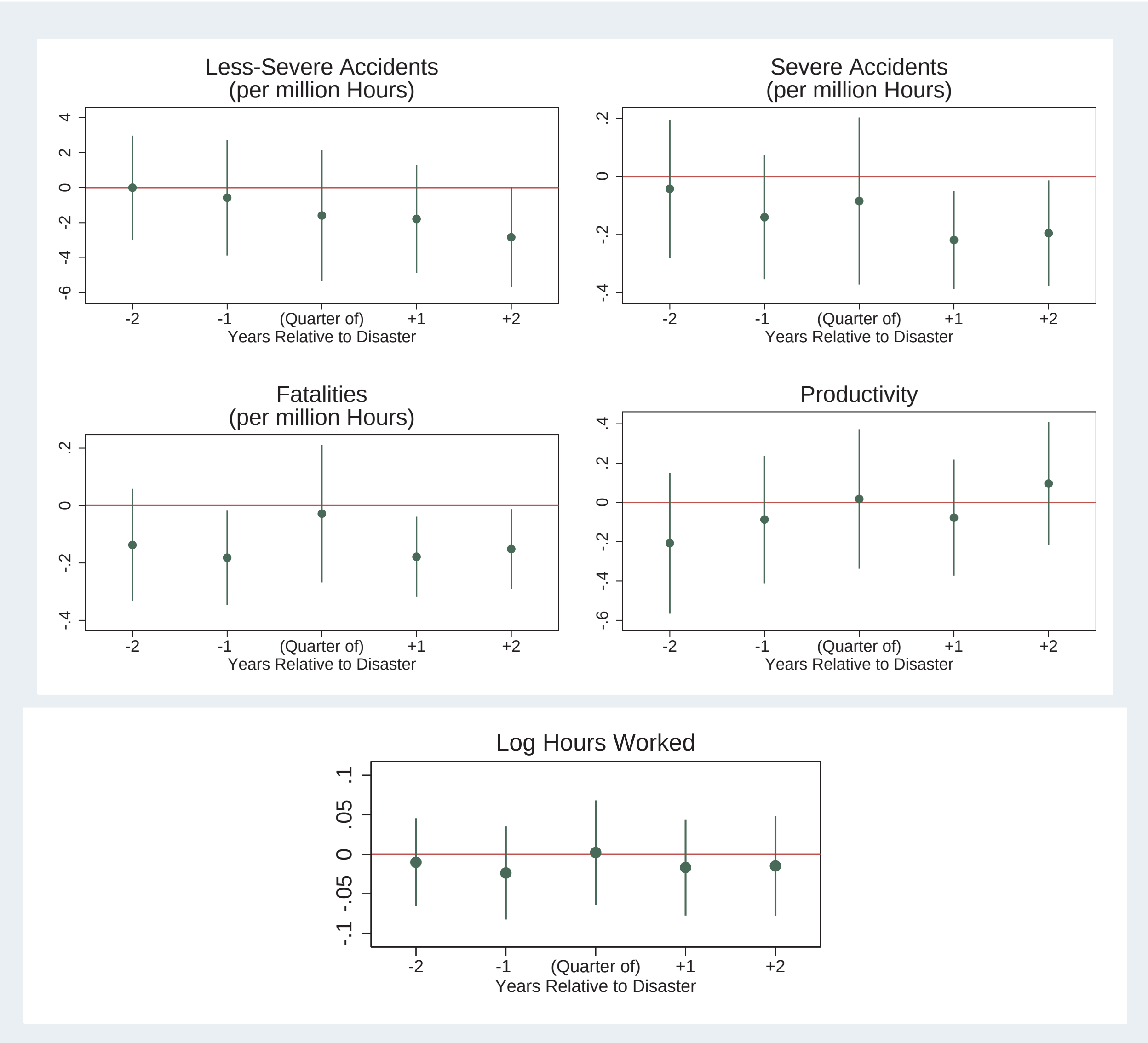

Each figure reports selected regressors from one regression. Dependent variables are indicated in the title. All regressions include mine fixed effects, quarterly fixed effects, and state dummies times hours worked and state dummies times employees (to control for economies of scale) as regressors and are weighted by mean hours worked at the mine. Standard errors are clustered at the mine level. Each line shows a 95% confidence interval.

Table A2: Regressions of Disasters Within 400 KM on Productivity and Safety

| | (1) Less-Severe Accidents (per million hours) | (2) Severe Accidents (per million hours) | (3) Fatalities (per million hours) | (4) Productivity | (5) Log Worker Hours |
|---|---|---|---|---|---|
| *Effect by Year* | | | | | |
| Two Years Prior to Disaster | -0.0063 | -0.043 | -0.14 | -0.21 | -0.010 |
| | (1.52) | (0.12) | (0.100) | (0.18) | (0.028) |
| One Year Prior | -0.58 | -0.14 | -0.18** | -0.087 | -0.024 |
| | (1.68) | (0.11) | (0.084) | (0.17) | (0.030) |
| Quarter of Disaster | -1.59 | -0.084 | -0.028 | 0.018 | 0.0022 |
| | (1.90) | (0.15) | (0.12) | (0.18) | (0.034) |
| One Year After | -1.78 | -0.22** | -0.18** | -0.078 | -0.017 |
| | (1.57) | (0.086) | (0.071) | (0.15) | (0.031) |
| Two Years After | -2.83* | -0.19** | -0.15** | 0.096 | -0.015 |
| | (1.46) | (0.092) | (0.071) | (0.16) | (0.032) |
| State× Hours Worked | Yes | Yes | Yes | Yes | No |
| State×Employees | Yes | Yes | Yes | Yes | No |
| N | 51,477 | 51,477 | 51,477 | 51,477 | 51,477 |
| $R^2$ Adj. | 0.04 | 0.00 | 0.00 | 0.31 | 0.03 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column reports the results of one regression. Dependent variable is indicated in the column title. All regressions include mine fixed effects and quarterly fixed effects as regressors and are weighted by mean hours worked at the mine. Standard errors, reported in parentheses, are clustered at the mine level.

Figure A2: Regressions of Disasters Within 400 KM on Productivity and Safety




Each figure reports selected regressors from one regression. Dependent variables are indicated in the title. All regressions include mine fixed effects, quarterly fixed effects, and state dummies times hours worked and state dummies times employees (to control for economies of scale) as regressors and are weighted by mean hours worked at the mine. Standard errors are clustered at the mine level. Each line shows a 95% confidence interval.

# Bibliography

Acuna, D. and Schrater, P. Bayesian modeling of human sequential decision-making on the multi-armed bandit problem. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, volume 100, pages 200–300. Washington, DC: Cognitive Science Society, 2008.

Aguiar, M. A. and Bils, M. Has consumption inequality mirrored income inequality? Technical report, National Bureau of Economic Research, 2011.

Alford, J. Can government regulate safety? The coal mine example. *American Political Science Review*, 74:745–756, 1980.

Amir, R. Supermodularity and complementarity in economics: an elementary survey. *Southern Economic Journal*, pages 636–660, 2005.

Attanasio, O., Battistin, E., and Ichimura, H. What really happened to consumption inequality in the US? Technical report, National Bureau of Economic Research, 2004.

Attanasio, O., Hurst, E., and Pistaferri, L. The evolution of income, consumption, and leisure inequality in the US, 1980-2010. Technical report, National Bureau of Economic Research, 2012.

Autor, D. H., Katz, L. F., and Kearney, M. S. Trends in US wage inequality: Revising the revisionists. *The Review of Economics and Statistics*, 90(2):300–323, 2008.

Battistin, E. Errors in survey reports of consumption expenditures. Technical report, IFS Working Papers, Institute for Fiscal Studies (IFS), 2003.

Becker, G. S. and Murphy, K. M. The upside of income inequality. *The American*, 1 (4), 2007.

Bergemann, D. and Valimaki, J. Bandit problems. 2006.

Berry, S., Levinsohn, J., and Pakes, A. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.

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

Bombardini, M. and Trebbi, F. Risk aversion and expected utility theory: an experiment with large and small stakes. *Journal of the European Economic Association*, 10(6):1348–1399, 2012.

Bresnahan, T. F. Measuring the spillovers from technical advance: mainframe computers in financial services. *The American Economic Review*, pages 742–755, 1986.

Brophy, J. and Hall, J. O. *A Miner's Life: An Autobiography*. University of Wisconsin Press, 1964.

Brown, K., Campbell, S. W., and Ling, R. Mobile phones bridging the digital divide for teens in the US? *Future Internet*, 3(2):144–158, 2011.

Brynjolfsson, E. The contribution of information technology to consumer welfare. *Information Systems Research*, 7(3):281–300, 1996.

Brynjolfsson, E., Hu, Y., and Smith, M. D. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science*, 49(11):1580–1596, 2003.

CBS. Nike will adjust ad that depicts West Virginia mine, 09 2010. URL `http://chicago.cbslocal.com/2010/09/02/nike-will-adjust-ad-that-depicts-west-virginia-mine/`. [Accessed: 2015 01 30].

Chetty, R. and Friedman, J. N. New evidence on the long-term impacts of tax credits. 2011.

Chetty, R. and Saez, E. Teaching the tax code: Earnings responses to an experiment with eitc recipients. *American Economic Journal: Applied Economics*, 5(1):1–31, 2013.

Chiappori, P.-A., Levitt, S., and Groseclose, T. Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*, pages 1138–1151, 2002.

Cutler, D. M. and Katz, L. F. Rising inequality? changes in the distribution of income and consumption in the 1980s. Technical report, National Bureau of Economic Research, 1992.

Dahl, G. and Lochner, L. The impact of family income on child achievement: Evidence from the earned income tax credit. Technical report, National Bureau of Economic Research, 2008.

Dalton, H. The measurement of the inequality of incomes. *The Economic Journal*, pages 348–361, 1920.

Deck, C., Lee, J., and Reyes, J. Risk attitudes in large stake gambles: evidence from a game show. *Applied Economics*, 40(1):41–52, 2008.

Dionne, S. Death payouts in BP spill start at $8 million, March 2011. URL http://www.wsj.com/articles/SB10001424052748703739204576228881749944162. [Accessed: 2015 01 30].

Ellison, N. B., Steinfield, C., and Lampe, C. The benefits of facebook friends: social capital and college students use of online social network sites. *Journal of Computer-Mediated Communication*, 12(4):1143–1168, 2007.

Energy Information Administration. *Quarterly Coal Report, April - June 2014*. U.S. Department of Energy, October 2014.

Energy Information Administration. *Annual Coal Report 2013*. U.S. Department of Energy, January 2015.

Erbas, S. N. and Sayers, C. L. *Is the United States CPI biased across income and age groups?* International Monetary Fund, 1998.

Everett, J. G. and Slocum, A. H. Cranium: device for improving crane productivity and safety. *Journal of Construction Engineering and Management*, 119(1):23–39, 1993.

Feenstra, R. C. and Shapiro, M. D. Introduction to" scanner data and price indexes". In *Scanner Data and Price Indexes*, pages 1–14. University of Chicago Press, 2003.

Fiedler, F. E., Bell Jr, C. H., Chemers, M. M., and Patrick, D. Increasing mine productivity and safety through management training and organization development: A comparative study. *Basic and Applied Social Psychology*, 5(1): 1–18, 1984.

Fishback, P. V. *Soft Coal, Hard Choices*. New York: Oxford University Press, 1992.

Fox, S. and Rainie, L. Part 1: How the internet has woven itself into american life. Technical report, Pew Research Center, 2014.

Fukuyama, F. *Political Order and Political Decay: From the Industrial Revolution to the Globalization of Democracy*. Farrar, Straus and Giroux, 2014. ISBN 978-0-374-22735-7.

Gittins, J. and Jones, D. Adynamic allocation index for the sequential allocation of experiments. *Progress in Statistics*, pages 241–66, 1974.

Goettler, R. L. and Gordon, B. R. Does AMD spur Intels to innovate more? *Journal of Political Economy*, 119(6):1141–1200, 2011.

Goldfarb, A. and Prince, J. Internet adoption and usage patterns are different: Implications for the digital divide. *Information Economics and Policy*, 20(1):2–15, 2008.

Goolsbee, A. and Klenow, P. J. Valuing consumer products by the time spent using them: An application to the internet. Technical report, National Bureau of Economic Research, 2006.

Gowrisankaran, G. and Rysman, M. Dynamics of consumer demand for new durable goods. *Journal of Political Economy*, 120:1173–1219, 2012.

Gray, W. B. The cost of regulation: Osha, epa and the productivity slowdown. *The American Economic Review*, pages 998–1006, 1987.

Greenwood, J. and Kopecky, K. A. Measuring the welfare gain from personal computers. *Economic Inquiry*, 51(1):336–347, 2013.

Gross, M. B., Hogarth, J. M., Schmeiser, M. D., et al. Use of financial services by the unbanked and underbanked and the potential for mobile financial services adoption. *Federal Reserve Bulletin*, 98(4):1–20, 2012.

Handbury, J. Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities. *Photocopy, Columbia University, New York*, 2012.

Hansen, L. P. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982.

Hausman, C. Corporate incentives and nuclear safety. *American Economic Journal: Economic Policy*, 6:178–206, 2014.

Hausman, J. Cellular telephone, new products, and the cpi. *Journal of business & economic statistics*, 17(2):188–194, 1999.

Hausman, J. Sources of bias and solutions to bias in the consumer price index. *Journal of Economic Perspectives*, pages 23–44, 2003.

Hausman, J. A., Pakes, A., and Rosston, G. L. Valuing the effect of regulation on new services in telecommunications. *Brookings papers on economic activity. Microeconomics*, pages 1–54, 1997.

Heathcote, J., Perri, F., and Violante, G. L. Unequal we stand: An empirical analysis of economic inequality in the united states, 1967–2006. *Review of Economic Dynamics*, 13(1):15–51, 2010.

Hendel, I. and Spiegel, Y. Small steps for workers, a giant leap for productivity. *American Economic Journal: Applied Economics*, 6(1):73–90, 2014.

Hensher, D., Daniels, R., and Battellino, H. Safety and productivity in the long distance trucking industry. In *Proceeding, 16th ARRB Conference, 9-13 November 1992, Perth, Western Australia; Vol 16, PART 4*, 1992.

Hobijn, B. and Lagakos, D. Inflation inequality in the United States. *Review of Income and Wealth*, 51(4):581–606, 2005.

Howard, P. N., Duffy, A., Freelon, D., Hussain, M., Mari, W., and Mazaid, M. Opening closed regimes: what was the role of social media during the arab spring? 2011.

Hoynes, H. W., Miller, D. L., and Simon, D. Income, the earned income tax credit, and infant health. Technical report, National Bureau of Economic Research, 2012.

Juhn, C., Murphy, K. M., and Pierce, B. Wage inequality and the rise in returns to skill. *Journal of political Economy*, pages 410–442, 1993.

Kniesner, T. J. and Leeth, J. D. Data mining mining data: MSHA enforcement efforts, underground coal mine safety, and new health policy implications. *Journal of Risk and Uncertainty*, 29(2):83–111, 2004.

Krueger, D. and Perri, F. Does income inequality lead to consumption inequality? evidence and theory. *The Review of Economic Studies*, 73(1):163–193, 2006.

Krueger, D., Perri, F., Pistaferri, L., and Violante, G. L. Cross-sectional facts for macroeconomists. *Review of Economic Dynamics*, 13(1):1–14, 2010.

Kuhn, P. and McAusland, C. Consumers and the brain drain: Product and process design and the gains from emigration. *Journal of International Economics*, 78(2): 287–291, 2009.

Laikin, M. *Lens design*. CRC Press, 2010.

Lewman, A. Journalists use tor to communicate more safely with whistle”blowers and dissidents. nongovernmental organizations (ngos) use tor to allow. *Advances in Cyber Security: Technology, Operation, and Experiences*, page 109, 2013.

Ley, E. Whose inflation? a characterization of the cpi plutocratic gap. *oxford economic Papers*, 57(4):634–646, 2005.

Marr, C., Charite, J., and Huang, C.-C. Earned income tax credit promotes work, encourages childrens success at school, research finds. *Washington: Center on Budget and Policy Priorities. http://www. cbpp. org/files/6-26-12tax. pdf*, 2013.

Melnikov, O. Demand for differentiated durable products: The case of the US computer printer market. *Economic Inquiry*, 51(2):1277–1298, 2013.

Meyer, B. D. and Sullivan, J. X. Measuring the well-being of the poor using income and consumption. Technical report, National Bureau of Economic Research, 2003.

Meyer, B. D. and Sullivan, J. X. Changes in the consumption, income, and well-being of single mother headed families. *The American Economic Review*, pages 2221–2241, 2008.

Meyer, B. D. and Sullivan, J. X. Consumption and income inequality in the US since the 1960s. *University of Chicago manuscript*, 2013.

Montorselli, N. B., Lombardini, C., Magagnotti, N., Marchi, E., Neri, F., Picchi, G., and Spinelli, R. Relating safety, productivity and company type for motor-manual logging operations in the italian alps. *Accident Analysis & Prevention*, 42(6): 2013–2017, 2010.

National Safety Council. *Injury Facts 2014 Edition*. National Safety Council, 2014.

Nordhaus, W. D. Do real-output and real-wage measures capture reality? the history of lighting suggests not. In *The economics of new goods*, pages 27–70. University of Chicago Press, 1996.

Obama, B. Remarks by the President on mine safety, 04 2010. URL `http://www.whitehouse.gov/the-press-office/remarks-president-mine-safety`. [Accessed: 2015 01 30].

Petrin, A. Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110(4):705–729, 2002.

Piketty, T. and Saez, E. Income inequality in the united states, 1913-1998 (series updated to 2000 available). Technical report, National bureau of economic research, 2001.

Pollak, R. A. Group cost-of-living indexes. *The American Economic Review*, pages 273–278, 1980.

Prais, S. J. Whose cost of living? *The Review of Economic Studies*, pages 126–134, 1959.

Rittenberg, L. and Manuel Jr, E. Sources of labor productivity variation in the US surface coal mining industry, 1960-1976. *The Energy Journal*, pages 87–100, 1987.

Rothschild, M. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9(2):185–202, 1974.

Ruggles, S., Alexander, J. T., Genadek, K., Goeken, R., Schroeder, M. B., and Sobek, M. Integrated public use microdata series: Version 5.0, 2010. URL https://usa.ipums.org/usa/cite.shtml. [Accessed: 2015 04 20].

Rust, J. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033, 1987.

Sawacha, E., Naoum, S., and Fong, D. Factors affecting safety performance on construction sites. *International Journal of Project Management*, 17(5):309–315, 1999.

Schmitz, J. A. What determines productivity? Lessons from the dramatic recovery of the us and canadian iron ore industries following their early 1980s crisis. *Journal of Political Economy*, 113(3):582–625, 2005.

Sider, H. Safety and productivity in underground coal mining. *The Review of Economics and Statistics*, pages 225–233, 1983.

Slesnick, D. T. *Consumption and social welfare: Living standards and their distribution in the United States*. Cambridge University Press, 2001.

Song, M. Measuring consumer welfare in the CPU market: an application of the pure-characteristics demand model. *The RAND Journal of Economics*, 38(2): 429–446, 2007.

Sutton, J., Palen, L., and Shklovski, I. Backchannels on the front lines: Emergent uses of social media in the 2007 southern california wildfires. In *Proceedings of the 5th International ISCRAM Conference*, pages 624–632. Washington, DC, 2008.

Topkis, D. M. Minimizing a submodular function on a lattice. *Operations research*, 26(2):305–321, 1978.

U.S. Bureau of Labor Statistics. May 2013 OES national industry-specific occupational employment and wage estimates, April 2014. URL http://www.bls.gov/oes/current/oes_ind.htm. [Accessed: 2015 02 10].

Viscusi, W. K. and Aldy, J. E. The value of a statistical life: a critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, 27(1): 5–76, 2003.

Walker, M. and Wooders, J. Minimax play at wimbledon. *American Economic Review*, pages 1521–1538, 2001.

Weber, R. et al. On the gittins index for multiarmed bandits. *The Annals of Applied Probability*, 2(4):1024–1033, 1992.

Weeks, J. L. and Fox, M. Fatality rates and regulatory policies in bituminous coal mining, United States, 1959-1981. *American Journal of Public Health*, 73(11): 1278–1280, 1983.