Essays on the Interaction Effects of Policies Across Jurisdictions

by

Ben J. Niu

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Economics) in the University of Michigan 2013

Doctoral Committee:

Professor James R. Hines Jr., Chair
Professor Alan V. Deardorff
Associate Professor Andrei A. Levchenko
Professor Joel B. Slemrod

DEDICATION

This dissertation is dedicated to my wife, Dana Marchessault Niu, for six years of patience, support, and health insurance.

ACKNOWLEDGMENT

To my committee members, thank you for all of your time and effort throughout this process. Your impact as motivators and resources has been immeasurable. To my wife, Dana Marchessault Niu, you moved to Michigan for me. What else can I say? To my parents, Li Niu and Li Zeng, thank you for always putting my education first. You gave up so much to make sure I had the best opportunities to succeed in life. To my friends and fellow graduate students, thank you for always being honest, for keeping me honest, and for being sources of amusement. Lastly, to my grandparents, thank you for supporting me unequivocally and unconditionally in every aspect of my life. My grandmothers, I will always remember you. My grandfathers, I know - I owe you two a boat.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ABSTRACT

Essays on the Interaction Effects of Policies Across Jurisdictions

by

Ben J. Niu

Chair: James R. Hines, Jr.

This dissertation analyzes the theoretical and empirical effects of cross-border policy dynamics, specifically state level bottle deposit-redemption systems and international corporate income taxation. As governments can only enact policies within their own jurisdictions, borders often become areas with policy discontinuities. Mobility of economic factors allows firms and consumers to take advantage/arbitrage across such divides. The presence of borders therefore carries important positive and normative policy implications.

Chapter 1 analyzes the effect of Michigan's deposit-redemption system on Michigan as well as Indiana and Ohio, two bordering states with no such policy. In conjunction with a sales tax differential between the three states, households are theoretically able to fraudulently redeem out-of-state bottles and evade use taxes via cross-border shopping. I hypothesize that evidence of this behavior should be reflected in the patterns of retail prices for

deposit eligible goods near these borders. Results from a simulated model of imperfect competition and mobile households suggest that the main determinant of prices is travel costs. I then empirically analyze an original dataset and find that retail prices are increasing and decreasing with distance from the border in Indiana/Ohio and Michigan, respectively. At the border, prices are generally higher on the Michigan side.

Chapters 2 and 3 analyze the strategic interactions between corporate income taxes of countries that compete to attract mobile firms and capital. Specifically, I analyze the equilibrium revenue implications when competing countries offer preferential tax rates for targeted tax bases. Chapter 2 uses numerical methods to generalize existing theoretical literature in analyzing a less restrictive model of international tax competition. In the context of bilateral tax agreements, I find that countries with greater productivity and population asymmetry generate greater equilibrium revenues by allowing for preferential policies. Conversely, more symmetric countries would do better by banning such policies. Chapter 3 focuses specifically on the equilibrium revenue effect of cross-country profit correlation, i.e., the degree to which firms generate the same profits across different countries. As profit correlation decreases, preferential regimes become more revenue-dominant.

CHAPTER I

Retail Bottle Pricing at the Border: Evidence of Fraudulent Redemptions and Use Tax Evasion

1.1 Introduction

In 1971, Oregon became the first state to pass legislation approving the enactment of a bottle deposit-redemption system. Colloquially known as "bottle bills", state-level deposit-redemption systems require consumers to pay an additional per-bottle fee on purchases of items such as sodas and beers sold in bottles and cans.[1] The deposit is then returned to the consumer when the empty bottles are redeemed at approved redemption centers, or via reverse vending machines (RVMs) located at larger retail locations. Because any individual retailer may have more or less deposits than redemptions, distributors and/or bottlers which operate on a larger geographic scale act as the financial intermediaries between retailers and the state. These distributors collect the deposits taken in by their associated retailers and then reimburse them for the outgoing redemptions. Originally, the bottle bill

[1]The usage of the word "bottles" refers generically to deposit-eligible beverage containers, and includes plastic and glass bottles as well as aluminum cans.

was meant to curb littering as cheap gas, roadway expansions, and changes in product packaging trends increased the amount of empty containers being thrown onto streets and highways. Within ten years of its enactment, the percentage of roadside litter attributable to soda and beer bottles decreased from 40% to 9%.[2] Since then, California, Connecticut, Hawaii, Iowa, Maine, Massachusetts, Michigan, New York, and Vermont have all adopted similar legislation.[3] More recently, however, the motivation for bottle bills has shifted from litter reduction to pro-recycling.

The deposit-redemption system was designed with the intent of being price neutral for consumers so as to not effect their consumption decisions. However, the economic and behavioral effects on consumers are far from neutral if we account for various redemption related costs, e.g., storage, travel, and time requirements. Additionally, the policy discontinuities created by these state-level laws generate arbitrage incentives for cross-border shopping. The bottle bill also incentivizes households to fraudulently redeem bottles, i.e., collect deposits on bottles for which none were paid. This is possible due to the fact that bottles are not state-specific. RVMs and redemption centers cannot distinguish between those purchased in Michigan, which paid the deposit, and those purchased out-of-state, which did not.

With the distinction of being the only state with a ten cent versus the common five cent deposit, Michigan is the most prominent example of such arbitrage. It also shares nearly 200 miles of land borders with its non-bottle bill neighbors, Indiana and Ohio. This bottle bill wedge creates incentives for fraudulent redemptions on both sides of the border. Michigan households can purchase in Indiana and avoid the deposit. If they also choose to fraudulently redeem, this gives them another ten cents per bottle. Likewise, Indiana

[2] Littering numbers are from a 2004 report published by the Oregon Department of Environmental Quality.

[3] Delaware had a bottle bill but replaced it in 2010 with a state-wide curbside recycling policy. The program is partially financed via a non-refundable deposit.

households can fraudulently redeem their purchases in Michigan.

Concurrently, a sales tax wedge incentivizes households in the opposite direction. Michigan levies a 6% sales tax while Indiana and Ohio have tax rates of 7% and 6.25%-7.75%, respectively. An even larger difference exists for soda as Michigan exempts soda purchases from all sales taxation. This tax wedge does two things. First, it dampens the effect of the bottle bill wedge by making it more expensive, net of the tax, to make purchases in Indiana and Ohio. Second, Indiana and Ohio households that make cross-border purchases at the lower Michigan tax rate are legally obligated to remit the tax difference to their home states. For example, an Indiana household that purchased $100 worth of beer in Michigan would have paid an additional $6 in sales tax. Because Indiana has a higher 7% rate and would have collected $7, the household owes the $1 difference in use taxes to Indiana. In practice, however, most households do not pay this difference and are in fact committing tax evasion.[4]

While the presence of the bottle bill and sales tax wedges generates incentivizes for cross-border purchases, fraudulent redemptions, and use tax evasion, no direct data exist to confirm and/or quantify these illegal activities. Therefore, this paper indirectly analyzes this hypothesis by studying the patterns of retail prices for deposit eligible goods near the borders between Michigan and Indiana/Ohio. If arbitrage incentives do have a significant impact on household behavior, then we would expect such effects to be capitalized in the retail prices. By simulating a theoretical model of mobile households, and empirically analyzing an original dataset of bottle prices, this paper looks to address two main questions. First, how do retail prices adjust as a function of distance to/from the border in these three states? Second, what is the magnitude of the border gap, assuming that one exists? As we move closer to the border, the strength of these wedges and their effects on retail prices should increase.

[4] Internet sales are also an area with significant use tax evasion. As a result, a number of states have been making a concerted push for increased support of the Streamlined Sales and Use Tax Agreement.

We would therefore expect to see a price-distance trend in the pre-tax, pre-deposit prices of bottled goods within each state. This would also suggest the presence of a sharp, price discontinuity at the borders themselves.

The results from the model suggest that the bottle bill and tax wedges are functions of the deposit amount, tax rates, and prices. More importantly, the relative sizes of each wedge are largely influenced by the level of household mobility. High levels of household mobility depress the overall level of retail prices in all states, which in turn decreases the relative effect of tax differences. Therefore, low travel costs result in positive border gaps and price-distance trends, i.e., Indiana/Ohio border prices are higher than Michigan border prices with prices in both states increasing over distance. Under high travel costs, the results flip. At intermediate levels of mobility, we see a combination of the two outcomes.

Empirically, I find that non-supermarket prices are generally increasing (decreasing) with distance to the border in Indiana and Ohio (Michigan), respectively, when taking into account other factors such as population, household income, and retail density. For example, I estimate that a two liter bottle of Coke purchased from an Indiana retailer located 30 minutes from the border is \$0.20 more expensive than one purchased at the border. Conversely, the price is \$0.36 cheaper for the same distance comparison on the Michigan side. Relative to the average price for two liter Cokes, this constitutes a difference of 12% and -21%, respectively. In regards to the border gaps, Michigan's non-supermarket (supermarket) border prices are generally higher (lower) than those on the Indiana and Ohio sides. For example, I estimate a border gap of -0.447 for the previous case. This implies that two liter Cokes on the Indiana border are approximately \$0.45 less than the same good sold adjacently on the Michigan border. However, this reverses in the case of 24 packs of Bud Lite at the Ohio-Michigan border where I estimate a positive border gap. Border retailers on the Ohio side are now

$1.04 more expensive than border retailers on the Michigan side. Given the estimated price patterns, this points to the existence of cross-border shopping, fraudulent redemptions, and use tax evasion. It also suggests that the effect of the bottle bill wedge is weaker relative to the effect of the tax wedge.

The remainder of the paper proceeds as follows. Section 2 provides further background on bottles bills, fraudulent redemptions, and relevant literature. Section 3 presents a theoretical model of household and firm behavior that incorporates these two wedges. A numerical simulation then generates predictions from the model. Section 4 describes the dataset and presents the results of the empirical analysis. Lastly, Section 5 concludes.

1.2 Background

1.2.1 Michigan's Bottle Bill

Bottle bills generally apply to beverages such as soda and beer that are sold in glass, plastic, and aluminum containers. Deposits are five cents in all bottle bill states except for Michigan, where it is ten cents, and California, where beverages larger than 24 fluid ounces are subject to a ten cent deposit. Maine and Vermont also levy a 15 cent deposit on stronger alcohol. Beverages sold out of the original container at venues are exempt from the deposit. In all bottle bill states, save for California, the actual deposits are exempt from taxation. This implies that the total purchase price of a $1.00 can of beer in Michigan is $1.16, with $0.06 coming from the 6% sales tax and $0.10 coming from the deposit. Indiana and Michigan have sales tax rates of 7% and 6%, respectively. Ohio has a state sales tax rate of 5.5% but the addition of local sales taxes pushes the total into the 6.25% to 7.75% range. Most counties, however, are in the 6.75% to 7% range.

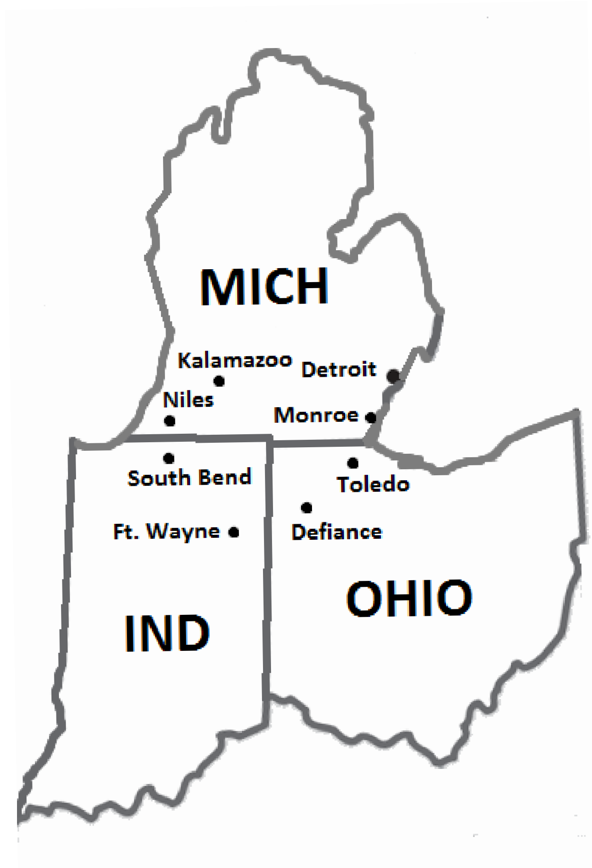
Bottles subject to a deposit are labeled as such near the UPC bar codes. The bar codes also serve to confirm the eligibility of certain goods when they are redeemed. Most bottle bill states allow for redemptions to occur via self-serve RVMs. RVMs are convenient because they are found at almost all large retail locations such as supermarkets and big box stores, and many times even at smaller retail locations such as pharmacies and convenience stores. These bar codes, however, do not allow RVMs to identify the state from which the bottle was purchased, i.e., bottles and bottle labels are identical across states. Therefore, RVMs are unable to identify those bottles for which a deposit was paid.

Michigan's bottle bill and its higher ten cent deposit are well known.[5] These fraudulent redemptions can be categorized into two types, although the law makes no distinction between the two. In-state fraud is committed when a Michigan household purchases bottles in Indiana or Ohio, where there is no bottle deposit, and redeems them in Michigan. Out-of-state fraud is committed by Indiana or Ohio households who bring their out-of-state bottles into Michigan. This does not, however, preclude Indiana households from redeeming bottles purchased in Michigan *if* deposits were collected. However, a redemption can only be claimed in the state from which it was purchased, i.e., a bottle purchased in Massachusetts cannot be redeemed in Vermont even though both states have five cent deposits.

Geographically, the length of the border and the distribution of population in the three states makes Michigan particularly vulnerable to cross-border shopping and fraudulent redemptions. In Michigan, there are two metropolitan statistical areas (MSAs) located near the southern border, Monroe and Niles-Benton Harbor. The Monroe MSA is located at the eastern-most side of the border (with Ohio) while the Niles-Benton Harbor MSA is located

[5] A plan to arbitrage between Michigan's ten cent deposit and New York's five cent deposit was the central plot in two episodes (1996) of the television show *Seinfeld*. In order to generate positive profits in the venture, the characters solved the travel cost dilemma by using a postal delivery truck.

Figure 1.2.1: Map of Midwest



on the western-most side of the border (with Indiana). Monroe benefits from being the intermediate point between the much more urban Detroit-Livonia-Wayne (MI) and Toledo (OH) MSAs. Comparatively, Niles is a smaller city surrounded by rural areas, and Benton Harbor leans heavily on Lake Michigan tourism. The more urban Michigan MSAs are located further north, aligned almost perfectly along the east-west running Interstate 94. As such, Michigan is generally characterized by a few smaller cities closer to the border with larger cities further away. In Indiana and Ohio, there are two major MSAs located right on the border, South Bend-Mishiwaka (IN) and Toledo (OH). Both are very urban areas with significant stretches of suburban areas as well. Also on the Indiana-Michigan border are the Elkhart-Goshen and Michigan City-La Porte MSAs. Elkhart and Goshen resemble smaller versions of South Bend while Michigan City is somewhat similar to Benton Harbor. It benefits from lake tourism but also acts as an intermediate manufacturing and shipping point to Chicago. Other urban MSAs such as Indianapolis (IN) and Columbus (OH) are much further south.

In all likelihood, some amount of fraud has always occurred in Michigan. It was not a significant issue until the 1990's and particularly in the 2000's when annual redemption rates consistently climbed above 94%.[6] Comparatively, other bottle bill states were more accustomed to seeing redemption rates in the 70%'s and 80%'s. While some of this difference is certainly attributable to the higher deposit, the presence of fraudulent redemptions was never more evident than in 1992 when Michigan's redemption rate was 100.41%. In that year, approximately 15 million additional bottles were redeemed. More recently in 2007, Michigan State Police made a number of arrests as part of *Operation Can Scam* in which a sophisticated network of individuals and business owners generated profits in excess of $500,000. This criminal ring took bottles from Toledo, Ohio and sold them to retailers in Detroit, Michigan who would then fraudulently redeem them in bulk. A 2009 state law, aimed at curbing fraudulent redemptions, mandated that bottlers put an identifying Michigan mark on bottles sold within the state. In 2011, the law was in the early stages of implementation when a federal court struck it down.[7]

The monetary effects of fraudulent redemptions impact the state as well as the whole-

[6] Redemption rates are from a report published by the Michigan Department of Treasury in 2010.

[7] The US Sixth Circuit Court of Appeals ruled in December 2011 that Michigan's RVM Antifraud Act violated the Commerce Clause in the case of *American Beverage Association v. Snyder et al.* By forcing an identifying mark on bottles sold in Michigan, it was creating an extraterritorial effect on out-of-state bottlers as well.

salers/bottlers. When redemption rates are less than 100%, the unclaimed deposits or escheat is split 75%/25% between the state of Michigan and the distributors. These state funds typically go towards program costs as well as other environmental initiatives. Because distributors aggregate the process across many retailers over a wide geographic area, they bear a significant hassle cost to facilitating the bottle bill. Even though there are provisions to reimburse distributors and retailers for these costs, such rates are fixed over time and usually range between only one to three cents per bottle. Therefore, fraudulent redemptions take money out of state coffers and introduce greater costs for distributors. Out-of-state fraud also directly removes money from the state economy.

1.2.2 Previous Literature

The theoretical literature regarding deposit-redemption systems is sparse. Dobbs [1991] analyzed a model which argued that in the case of littering, the optimal Pigouvian policy includes both a consumption/disposal tax and an user fee. Intuitively, the consumption of a good generates a number of social costs stemming from production externalities, littering (eyesore and clean-up costs), and proper disposal (landfill, recycling process costs). Individually, a consumption tax does not distinguish between littering and non-littering costs because it is set at the marginal, total social cost. A consumer that does not litter therefore still pays for the associated littering costs when it should in actuality only be assessed on the proper disposal costs. Alternatively, a user fee for proper disposal does nothing to discourage littering. Dobbs showed that having both taxes, with the user fee being a subsidy that refunds the marginal social costs associated with littering, i.e., a deposit-redemption system, generates the greatest welfare gains. Eggert and Weichenrieder [2004] also analyzed the optimal mechanism question. Their paper claimed that deposit-refund systems are never optimal unless additional taxes are included to extract the surplus gained by producers. The surplus is gained due to the assumption that producers can adjust prices in response to a bottle bill, and claim some fraction of the escheat. The two papers differ primarily in regards to assumptions on the level of market power and competition.[8]

More recent papers have empirically studied the behavioral aspects of bottle bills. Viscusi et al. [2013] used survey data from a national sample of over 3000 bottled water consumers to analyze how individual characteristics influence the likelihood to recycle with and without a bottle deposit. They found that most households follow an all or nothing recycling pattern with 45% always/fully recycling and 25% never recycling. After including a five cent bottle deposit, these proportions changed to 62% and 8% for redeemers and non-redeemers, respectively. From their survey data, they also found that the effect of a bottle deposit is strongest for those with lower incomes. This is confirmed by Ashenmiller [2011] who found that low income individuals derive a non-trivial amount of income via the bottle-redemption system. Using responses from a survey of redeemers in Santa Barbara, CA, she estimated that households earning less than $10,000 generated $340 annually, and $428 when limiting the sample to Spanish-speaking households. Intuitively, the bottle bill acts as a mechanism that sorts labor such that those with low wages efficiently redeem both their own and potentially other consumer's bottles. Ashenmiller [2010] analyzed a possible externality associated with the supplemental income granted to low income individuals via bottle bills. Specifically, she used the state-specific timing of bottle bill roll-outs to compare city-level petty crime rates in states with and without bottle bills. On average, crime rates were 11% lower in states with bottle bills which the author argues is due to the dampening effect of this added income on incentives for criminal behavior.

[8]Additional papers on the topic include Fullerton and Wu [1998] and Fullerton and Wolverton [2000].

Other papers have analyzed similar instances of cross-border arbitrage and smuggling. Bhagwati and Hansen [1973] analyzed the incentives to circumvent tariff and tax laws, as well as the welfare implications of such import/export smuggling. As with the topic discussed in this paper, economic data that directly quantifies these illegal activities are either missing or noisy. For example, soft smuggling techniques affect official export and import statistics by under-reporting or mis-categorizing goods. Lovenheim [2008] examined the likelihood of smokers to cross-border shop at lower priced/taxed jurisdictions using a survey of smokers' behavior. He estimated that up to 25% of smokers cross-border shopped or purchased in border locations because of price differences. Merriman [2010] also analyzed the degree of cross-border shopping in the cigarette market. Using littering data in the highly tax stratified Chicago area, he found that a one mile increase in distance to a low tax border increased the probability of finding a littered "home" cigarette pack by one percent. Additionally, there is a sizable border effects literature. The seminal paper, Engel and Rogers [1996], found that price variation within countries is far lower than he price variation across countries, even in the case of similar and close neighbors. They estimated that the US-Canadian border generates a price differential equivalent to 75,000 miles of within-country distance.

To my knowledge there have not been any academic papers specifically studying the effect of bottle bills and fraudulent redemptions on the spatial relationship between prices and border.[9] There has been, however, work on cross-border shopping and its relationship with optimal local tax rates. Kanbur and Keen [1993] and Agrawal [2012] posited that local jurisdictions adjust their local sales tax rates to account for differences in state sales tax

[9] One report in the late 1990's, *Analysis of Foreign Containers in the Michigan Deposit Stream*, was produced by a consulting group for the Michigan Beer and Wine Wholesalers Association in their push to enact stricter anti-fraud legislation. This report used surveys of retailers to estimate that roughly $16 million in redemptions came via fraud. A separate report commissioned by the state suggests that this number is conflated due to biased sampling in a small region of the state.

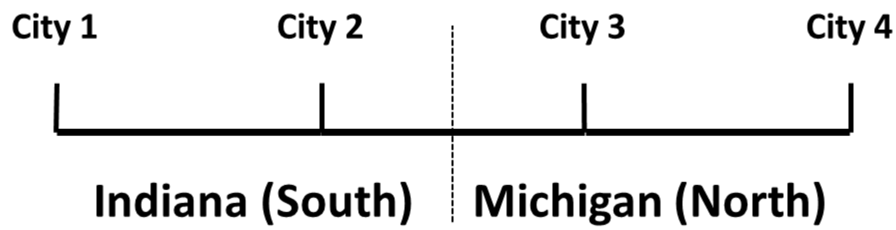
rates when cross-border shopping is present. Whereas this paper hypothesizes that arbitrage incentives would be capitalized in retail prices, this local sales tax literature hypothesizes that it is capitalized in these local taxes. However, it is important to note that not all states allow for separate local sales taxes including two of the three states studied in this paper.

1.3 Model

To model the behavior of households and firms, assume that the population is divided amongst two states and four cities. Along the lines of a Hottelling model, the four equally spaced cities are situated along a straight line such that Cities 1 and 2 are located in the southern state, Indiana, while Cities 3 and 4 are located in the northern state, Michigan. The distance between any two consecutively numbered cities is equal to one unit of distance. The relevant variables in the model, all denoted in cents, are V_i (value of consumption), c_i^T (marginal cost of travel), and c_i^R (marginal hassle cost of redeeming) for household i. Households are heterogeneous in the sense that these three positive variables are distributed according to F_V, F_T, and F_R, respectively. Cities, however, are identical in the sense that the distributions are not city specific. Additionally, all households receive new variable draws each period.

In each period, households choose where to purchase the (non-durable) bottled good and whether to redeem after observing the prices. Firms set profit maximizing prices in each of the four cities, taking into account the mobility of households. This occurs in the context of an infinitely repeated game.

Figure 1.3.1: Hottelling Line



1.3.1 Households' Problem

Households derive benefit V_i from the consumption of a homogenous, single-bottle good, e.g., a two liter bottle of Coke, which lasts for the duration of a period. Expressed as the monetary willingness to pay, V_i is household specific but not city specific. While households are unable to change their city of residency, they can travel between cities at a marginal cost of c_i^T cents per unit of distance (round trip). A household residing in City 1 therefore incurs a travel cost of $|j-1|c_i^T$ if it chooses to purchase in city j. These mobile consumers will decide where to purchase the good by maximizing (minimizing) the total benefit (cost) associated with each location. Let p_j and τ_j denote the retail price and sales tax rate, respectively, in city j. In the cases of Indiana and Michigan, there are no local sales taxes so $\tau_1 = \tau_2 \equiv \tau_{Ind}$ and $\tau_3 = \tau_4 \equiv \tau_{Mich}$. To simplify on notation, let $\hat{p}_j$ be defined as the net-of-tax price in city j. Due to Michigan's bottle bill, all purchases made in Cities 3 and 4 are subject to the deposit. As the good is comprised of a single bottle, ten cents is added to the net-of-tax prices. Households do not have to purchase the good if the total benefit drops below zero.

Households must also decide whether or not to redeem the good. The benefit to redeeming is the return value of the deposit. The cost of redeeming is captured by c_i^R.[10] Let c_i^R be the

[10] Assume that fraudulent redemptions do not result in any additional/special costs relative to legal redemptions.

marginal hassle cost of redeeming a single bottle stemming from storage and RVM usage. For households in Indiana that are not purchasing in Michigan, redeeming requires additional travel across the border and thus an additional travel cost. For households in Indiana that choose to purchase the good in Michigan, they are already incurring the travel cost so the cost of redeeming is limited solely to the hassle cost. For households in Michigan, I assume that the travel cost of redeeming is zero.[11]

As an example, consider the incentives of a household residing in Indiana.[12] More specifically, household i in City 2 is contemplating purchasing at home or in City 3. If the household chooses to not redeem, it will receive (1.3.1) and (1.3.2) in total benefit by purchasing at home or purchasing in City 3, respectively.

$$V_i - (1 + \tau_{Ind})p_2 \tag{1.3.1}$$

$$V_i - (1 + \tau_{Mich})p_3 - 10 - c_i^T \tag{1.3.2}$$

Absent redemptions, the household will purchase in the other cities and/or state if the potential net-of-tax, net-of-deposit price savings exceed the travel costs. As would be expected, households with lower marginal travel costs are more likely to purchase away from home. If the household decides to redeem, its potential total benefit is amended by the net redemption term, $10 - c_i^R$, plus an additional travel unit of travel cost, $-c_i^T$, if it was originally purchasing at home. Unless an Indiana household is making purchases in City 4, assume that all redemptions are made in the closest border city, City 3.

[11] Alternatively, RVMs are commonly found at retailers with the relatively largest sales and most customers, we can assume that travel costs associated with redemptions are already captured by the purchase decision.

[12] The full set of household equations is given in Appendix A.

$$V_i - (1 + \tau_{Ind})p_2 + 10 - c_i^R - c_i^T \tag{1.3.3}$$

$$V_i - (1 + \tau_{Mich})p_3 - c_i^R - c_i^T \tag{1.3.4}$$

Comparing (1.3.3) and (1.3.4), we can see that the only difference lies in the net-of-tax, net-of-deposit price. Given the 7% and 6% sales tax rates in the two states, City 2 can support a higher price difference of at most:

$$p_2 - p_3 \leq \frac{10 - 0.01p_2}{1.06} \tag{1.3.5}$$

and still be cheaper on net for its residents than City 3. For a per bottle price of 100 cents on the Michigan border, this implies that the Indiana price can be approximately 8 cents more. Fraudulent redemptions therefore have an obvious effect on the choice of purchase location for those redeeming households. For non-redeeming households, the potential up-charge is even greater since households would need to invest an additional unit of travel cost to purchase in Michigan, i.e., this adds an additional c_i^T term to the numerator in (1.3.5).

For households residing in Michigan, the purchase-redemption choice is similar. All else being equal, households have an incentive to purchase in Indiana due to the absence of the deposit. To attract redeeming households in City 3, City 2 cannot support as high of a price difference since the benefit of not paying the deposit is mitigated by the extra travel cost.

$$p_2 - p_3 \leq \frac{10 - 0.01p_2 - c_i^T}{1.06} \tag{1.3.6}$$

Unlike in Indiana, notice that the redemption choice for Michigan households is independent of the location choice, i.e., (1.3.6) is the same for redeemers and non-redeemers alike. For households in Indiana, the redemption choice is tied to the purchase location as location determines whether or not a household incurs an additional travel cost. Households in Michigan always redeem at home with no additional travel costs. What this implies is that the ability to fraudulently redeem does not induce previously non-redeeming Michigan households to redeem. Only redeeming Michigan households will be affected. Indiana households, on the other hand, can be induced to purchase more at home and redeem.

To account for goods with multiple bottles/deposits, we can scale up the single bottle analysis if we assume that the marginals, V_i, c_i^T, and c_i^R, are constant over different quantities. Note that c_i^T becomes both the marginal and average travel cost associated with an additional unit of bottle-distance. p_j in this case would represent the average price of a bottle. Given that the number of bottles per good is fixed, these assumptions imply that a household will either buy or not buy the good, which avoids issues with partial purchases.

1.3.2 Firms' Problem

On the firm side of the model, assume that there are a finite and fixed number of identical, zero cost retailers which sell the bottled good.[13] Within each city, these retailers collude/cooperate to set harmonized *city-specific* prices, $p_j \geq 0$, where prices are denoted in one cent increments. Imperfect competition is crucial in this setting. The pre-tax, pre-deposit prices would be identical across cities in a model of perfect competition. In the context of this infinitely repeated game, assume that cities set their prices simultaneously

[13]This implicitly assumes that production/wholesale costs are the same across cities and states. This is a simplification as there are differences. Per gallon of beer, Indiana, Michigan, and Ohio levy \$0.12, \$0.20, and \$0.18 in additional excise taxes applied at the bottler/wholesaler level. For a six pack of 12 ounce beer, this corresponds with state excise taxes of approximately \$0.07, \$0.11, and \$0.10, respectively. Figures are from the Tax Foundation.

each period. They can not, however, collude across cities or across different periods. I am, in essence, analyzing a four firm oligopoly model of Bertrand competition. The level of analysis will therefore focus on the single period Nash Equilibrium prices across the four cities, and the corresponding consumer behavior in the single period game. Each city's revenue maximizing price will be a function of the other cities' endogenous prices and the exogenous sales tax rates. The prices will also depend on the characteristics of the households in each city as given by the three main variables. Assume that retailers set their prices prior to the realizations of V_i, c_i^T, and c_i^R. Households make their purchase-redemption decisions under full information while retailers know only of the distributions.

What would we expect the equilibrium prices to look like *a priori*? Consider a scenario where there is no bottle bill and there are a large number of cities located beyond Cities 1 and 4. Far from the border, there would exist a prevailing equilibrium price, $\tilde{p}$, across the cities. With a tax difference at the border, Cities 2 and 3 would charge different prices with this effect dissipating with distance from the border. Prices on either side would asymptote towards $\tilde{p}$ since higher (lower) border prices would pull up (down) prices in neighboring cities. With the introduction of a bottle bill and fraudulent redemptions, p_2 (p_3) should rise (fall). Again, the effect should dissipate away from the border. Overall, the direction of the border discontinuity depends on the relative size of the tax wedge and the bottle bill wedge. For cities far away from the border, there should be little to no effect.

1.3.3 Simulation Setup

To analyze the equilibrium price implications of this four city model, I numerically simulated the model under a range of plausible parameter values.[14] Each city has a population of 5000

[14] The model is solved numerically, as opposed to analytically, due to the severe tractability issues stemming from the number of inequalities/cases and non-convexity of costs. The optimal local sales tax model in

households. For each household, I draw values for V_i, c_i^T, and c_i^R from their respective distributions. F_v, F_T, and F_R are assumed to be uniformly distributed and bounded from below at zero. The maximum value of the good and willingness to pay, $\bar{V}$ is fixed at 500 cents.[15] For the marginal travel costs, I varied the upper bounds between $\bar{c}^T \in \{100, 150, 200\}$. This range of moving costs typically generated intercity movements between 0% and 30%. Upper bounds on the hassle cost of redemption were set at $\bar{c}^R \in \{11, 12, 13\}$. These numbers were selected to generate a realistic range of redemption rates, 77%-91%, for Michigan households that purchase the good.

Following the actual tax rates, τ was set at 7% for Cities 1 and 2, and 6% for Cities 3 and 4. The sales tax rates in Ohio range between 6.25%-7.75% due to the presence of local sales taxes in addition to the state sales tax. On average, the combined sales tax rate is approximately 7%. I also simulate a 7% and 0% sales tax scenario to model the case of soda as it is exempt from taxation in Michigan. To break potential ties in cases where households are indifferent between various scenarios, I introduced an additional term, θ, which represents the ratio between the travel costs for purchases versus redeeming.[16] In the simulations, I assumed a θ of 1.01. Given a set of prices, households calculate the total benefit for each purchase-redeem combination and select the most advantageous one. Multiplying the number of households purchasing in city j by p_j gives us the amount of revenue for city j's retailers.

For the retailers, the choice of p_j depends on the characteristics of the households as well as the prices in other cities. Recall that all firms within a given city collude but do

Agrawal [2012] gets around this issue by limiting consumers to purchase at most one city away which drastically decreases the state space.

[15] This was selected ex post to generate equilibrium prices in the range of actual prices.

[16] For example, a household in City 2 would face the same travel costs if it purchased in City 1 and redeemed in City 3, as it would if it purchased and redeemed in City 4.

not collude across cities or periods. We are therefore interested in identifying the single period Nash Equilibrium quartet of cities' prices. To calculate the equilibrium prices, I used an iteration method of best responses. Starting from a set of harmonized prices, each city took turns re-optimizing their prices in response to the prices in the other cities. Moving in sequential order, this continued until they reached a stable Nash/Bertrand Equilibrium set of prices. Searching for an equilibrium in this manner guarantees stability and rules out potentially unrealistic and unstable equilibria. These equilibria would also be trembling hand perfect equilibria.

In a number of cases, however, existence was an issue, i.e., the iteration process failed to converge. Instead, cities' best responses looped or cycled amongst a closed set of prices. To increase the likelihood of convergence, I included in the iteration process a 1% threshold parameter. The threshold creates a range of inaction, i.e., sS policy, such that cities only re-optimized if the new price generated a revenue increase equaling at least 1% of the current revenues. We can interpret this threshold as an adjustment cost for changing the price. Alternatively, it can reflect uncertainty/variability in the characteristics of the population in the current period. Because retailers only know of the distributional characteristics of each household, they may be loathe to make changes to the price unless there is sufficient confidence of a positive revenue gain. The benefit to including a threshold parameter is that it increases the likelihood of convergence. The tradeoff is that it also enlarges the set of possible equilibria.[17] A threshold of 1% was selected to balance between these two tradeoffs. Additionally, I ran the iteration process with a range of different starting prices between 100 and 200 cents. Taking the average of such results provided a more informative picture of the set of likely equilibrium prices and revenues.

[17] For every stable equilibrium, there may exist other Nash Equilibria where prices are only a cent off and within the range of inaction.

A potential issue with the four city model/setup lies in the fact that the outlying cities, Cities 1 and 4, only face one city of contiguous competition whereas the inner cities face two. This would allow Cities 1 and 4 to always support higher prices, and generate a price-distance trend independent of the two wedges. One obvious solution would be to simply add another set of cities beyond the end cities. With enough cities, the pulling up effect would theoretically dissipate such that there would be negligible effects for those relevant cities close to the border. Computationally, additional cities presents significant challenges given the increases to the choice and state spaces.

Instead, I approximated this larger multi-city model by fixing the prices in Cities 1 and 4 such that only the two border cities can optimize their prices. Implicitly, this assumes that the outlying cities are sufficiently far away from the border and unaffected by the various border effects. The challenge lies in determining the appropriate fixed price that does not artificially generate a price trend by assuming a very low or very high fixed price. The correct price in this case is the prevailing equilibrium price, $\tilde{p}$, that would have occurred in the multi-city model without the bottle bill for cities far from the border. To calculate $\tilde{p}$, I simulate the four city model without the bottle bill while fixing p_1 and p_4 at some candidate p and allowing p_2 and p_3 to react. If the candidate price is less than the true $\tilde{p}$, then the resulting equilibrium p_2 will be higher than p. If the candidate price is higher than the true $\tilde{p}$, then p_2 will be lower than p. The correct p that most closely approximates $\tilde{p}$ is the first price after this flip occurs. Denote the fixed price p as p^*.

I therefore calculate p^* for each of the different levels of households mobility. With different tax rates in each state, 7% and 6%, there are two $\tilde{p}$ terms - one for each state - and thus two separate fixed p^* terms for p_1 and p_4. I calculate these by assuming that all four cities have the same tax rate. Fixing p_1 and p_4, the correct p^* is the price where the equilibrium p_2 and p_3 are equal to p^*. For the cases of 7%, 6%, and even 0% tax rates, the calculated p^* terms were all equal. This is likely due to the 1% threshold and, to a lesser extent, the discrete prices. Only in one case were the p^* terms different. Under a tax rate of 0% and $\bar{c}^T = 200$, p^* was equal to 91 cents as opposed to 90 cents for the 7% and 6% cases. For each of the tax and $\bar{c}^T$ cases, I therefore simulate the model fixing p_1 and p_4 at the p^* which corresponds with the given distributional parameters. Notice that both the prevailing price $\tilde{p}$ and p^* are decreasing with household mobility.

1.3.4 Simulation Results

Define the border gap as the difference in price between City 2 and City 3, $p_2 - p_3$. The price-distance trends in Indiana and Michigan are given by $p_1 - p_2$ and $p_4 - p_3$, respectively. Table 1.1 presents the resulting equilibrium prices (in cents) when taxes rates are set at 7% and 6%. The maximum hassle cost of redemption, $\bar{c}^R$, is set at 12 cents while the deposit is ten cents.[18]

Table 1.1: Simulated Prices in Cents - 7% and 6% Taxes with $\bar{c}^R = 12$

$\bar{c}^T$	p_1	p_2	p_3	p_4	Border Gap	IN Trend	MI Trend
100	55	47	67	55	-20	8	-12
150	74	72	102	74	-30	2	-28
200	90	92	134	90	-42	-2	-44

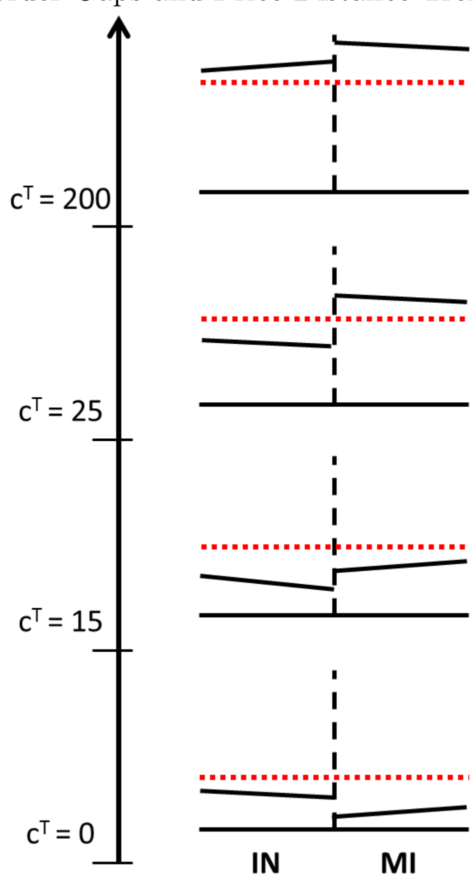
Figure 1.3.2 graphically presents a general representation of these results. Note that the pictures are not drawn to scale. The signs of the border gaps and price-distance trends are the main points of focus. The horizontal dotted line represents the prevailing price levels for

[18]The prices shown are averages of the equilibrium prices calculated under the different starting points.

those hypothetical cities located far from the border. Recall that p_1 and p_4 are fixed at the $\tilde{p}$.

In regards to the border gap, I find that the simulated equilibrium prices on the Indiana border are lower than prices on the Michigan border in the three $\bar{c}^T$ cases considered. In regards to the price-distance trends, the Indiana trend flips from being positive to negative when $\bar{c}^T$ increases to 200. The Michigan trend is negative in this entire range. Overall, the biggest determinant of the border gap and price-distance trends appeared to be the level of household travel costs.

Figure 1.3.2: Border Gaps and Price-Distance Trends for Various $\bar{c}^T$



Travel costs have a two-fold impact on the border gap. High travel costs depress the bottle bill wedge by making it difficult for Indiana households to travel for redemptions. More importantly, high travel costs increase the overall price levels in the cities. If households are less mobile, then there is less intercity competition which allows for greater monopoly power and a higher prevailing price level. The tax wedge increases in strength relative to the bottle bill wedge as prices rise. This gives Michigan retailers an advantage over Indiana retailers and results in a negative border gap. Conversely, lower travel costs push down overall price levels. At lower prices, the effect of the tax wedge decreases relative to the bottle bill wedge. When re-simulating under lower values of the travel cost parameter, I now find that the border gap becomes positive for $\bar{c}^T$ less than 15 as seen in Figure 1.3.2. At these lower travel costs, we have the situation where not only is the travel cost for redemption smaller, but so is the relative cost of buying in Indiana at a higher tax rate. Notice, however, that the border gap when $\bar{c}^T = 100$ is negative even though a 1% tax difference is smaller than the deposit at the calculated equilibrium prices. This indicates that additional factors, e.g., intercity competition and stickiness of prices, play a role such that prices do not perfectly counteract the net price difference.

From Table 1.3.2, we can see that Indiana's price-distance trend flips at $\bar{c}^T = 200$. Michigan's trend is negative in all three cases. Re-simulating under lower levels of travel costs, I find that Michigan also flips from negative to positive for $\bar{c}^T \leq 25$. As discussed previously, a low $\bar{c}^T$ means that households are very mobile which increases the degree of price competition. Because competition is more pronounced at the border, prices for Cities 2 and 3 fall below the prevailing at very low values of $\bar{c}^T$. Conversely, as $\bar{c}^T$ increases and the population as a whole becomes less mobile, intercity competition decreases and $\tilde{p}$ increases as is evident from Table 1.3.2. Border prices are also increasing but they increase at a faster rate. This causes the price-trends to flip and become negative. Intuitively, when households become more immobile, City 3 now has an advantage by being near City 2 as it faces weaker competition due to the bottle bill wedge. This allows City 3 to attract cross-border shoppers *and* keep their domestic consumers. Michigan's trend therefore flips from negative to positive as $\bar{c}^T$ increases. The higher levels of p_3 also pull up p_2 until both are above $\tilde{p}$ although this occurs at a relatively higher $\bar{c}^T$.

Table 1.2: Simulation Results Summary

Border Gap:	$p_2 - p_3$	Positive for $\bar{c}^T \leq 15$ Gap increases with deposit Gap decreases with $\tau_{Ind} - \tau_{Mich}$
Indiana Trend:	$p_1 - p_2$	Positive for $\bar{c}^T \leq 200$ Slope decreases with deposit Slope increases (?) with $\tau_{Ind} - \tau_{Mich}$
Michigan Trend:	$p_4 - p_3$	Positive for $\bar{c}^T \leq 25$ Slope increases with deposit Slope decreases with $\tau_{Ind} - \tau_{Mich}$

Table 1.2 summarizes the effect of $\bar{c}^T$ on the border gap and price-distance trends. It also summarizes the effect of different deposits and tax differentials. Switching from a ten cent to an eight cent deposit decreases the border gap by approximately a cent under the three previous mobility cases. As expected, a decrease in the deposit also increases (decreases) the price-distance trend in Indiana (Michigan). Moving from the 7% and 6% case to the 7% and 0% case decreases the border gap by an average of five cents in the three cases. In the simulations, the Michigan price-distance trend decreases for an increase in the gradient. On the Indiana side, the price-distance trend increases with the gradient except in the case

of $\bar{c}^T = 200$. This is due to the fact that the tax rates also affect prevailing price levels in each state. Since we now have a direct price effect as well as an indirect tax competition effect, there is less of a clear cut pattern. Most importantly, however, the level of household mobility is still the main determinant of the price-distance trend.[19]

1.4 Empirical Analysis

The theoretical model generates a number of implications. Central to this paper is the relationship between the states' border prices as well as the correlation between intrastate prices and distance from the border. To empirically analyze the actual price-distance trends, the basic strategy is to regress the different pre-tax, pre-deposit retail prices of a bottled good on the distance from the border for each corresponding retailer location. This regression can be run separately for each state sub-sample, or over the entire sample but with state interactions and fixed effects. For the border gaps, a similar approach can be used on the Indiana-Michigan and Ohio-Michigan sub-samples. The coefficients of interest those on the distance variable and state dummies. Below, I describe the dataset and the empirical methodology.

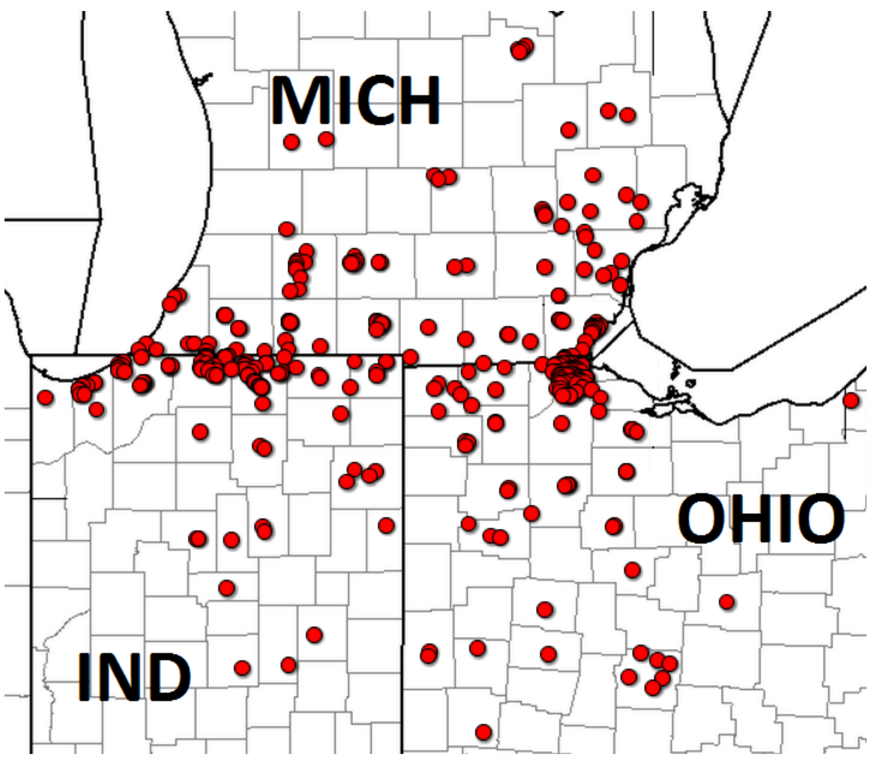
1.4.1 Original Dataset of Retail Prices

Using the Google Maps and Yellow Pages websites, I created a list of convenience stores, gas stations, grocery stores, liquor stores, pharmacies, and supermarkets within 150 miles of

[19] Given the model, it would be difficult to produce a reasonably justifiable estimate of the *levels* of cross-border shopping, use tax evasion, and fraudulent redemptions that are occurring. While the empirical analysis can motivate a range of household mobility with which to calibrate the model, the assumptions needed to make such an estimation tractable outweigh the accuracy of such an estimate.

the border.[20] A smaller subset of specific stores was then selected to identify those retailers from whom prices would be collected. Sampling wise, greater emphasis was placed on areas closer to the borders and on retailers with relatively greater bottle sales, i.e., liquor stores, grocery stores, and supermarkets. Otherwise, the selection of retailers was random.[21] To get a measure of distance to the border, I used Google Maps to calculate the routes with the shortest driving times. As opposed to a straight line, "as the crow flies" distance, Google Maps' driving directions are a more realistic measure of effective distance. For each location, I recorded the driving time in minutes and the corresponding distance in miles.[22]

Figure 1.4.1: Map of Retailers Sampled



[20]Existing UPC/scanner datasets focus predominantly on supermarket prices.

[21]See Appendix B for more details.

[22]Distance is calculated to the fastest *drivable* point on the border. Given the lack of available information to systematically generate the most likely, nearest cross-border retailer, this is a second best alternative.

The collection of retail prices was carried out via a combination of phone surveys and in-store visits conducted between December 2012 and April 2013. At each location, the following bottle pricing data were collected: the pre-tax and pre-deposit prices for a two liter bottle of Coke, 12 pack of Coke in cans, six pack of Budweiser in bottles, and 24 pack of Bud Lite in cans. Aside from the two liter Cokes, all other cans and bottles were of the standard 12 ounce variety. In Ohio, smaller retailers sometimes charged a higher price for cold versus warm 24 packs of Bud Lite, typically a dollar extra. In these cases, I collected prices for the warm version. For each price observation, I noted if the given price was marked as being on sale. Additional demographic and economic controls were obtained from the US Census, specifically the American Community Survey (2011 - 5 Year) and the County Business Patterns (2010). All of the Census variables were specified at the standardized ZIP code level (ZCTA5). Analogous data at the city, town, and township levels were taken from the website www.city-data.com.

In total, the dataset included 346 retailers. Driving distances ranged between 0 and 150 miles from the border while driving times ranged between 0 and 165 minutes. Column 5 in Table 1.3 shows the number of retailers sampled by state. It also provides a breakdown of the number of price observations for the four bottled goods. Due to the fact that not all retailers carry the same sets of goods, there are fewer price observations than retailer observations. In the case of soda, some stores did not stock any quantities larger than six packs while others only sold Pepsi products. In the case of beer, a number of stores, particularly pharmacies, either did not sell Bud Lite in quantities larger than 15 packs or did not sell any alcohol period. Additionally, the number of price observations denoted as being on sale is given in parentheses.

Table 1.3: Count of Observations (# Sales)

	2 Liter Coke	12 Pack Coke	6 Pack Bud	24 Pack BudLite	Retail Obs
Indiana	99 (29)	92 (26)	49 (2)	70 (25)	110
Michigan	108 (30)	91 (31)	97 (0)	105 (24)	126
Ohio	86 (20)	81 (22)	77 (0)	91 (12)	110
Total	293 (79)	264 (79)	223 (2)	266 (61)	346

The choice of these specific goods was purposely made to address a number of key points. I selected major brands in commonly sold quantities to facilitate greater comparability by mitigating potential issues with missing observations. As is evident, the bundle varied along the two dimensions of quantity and beverage type. By selecting four products with different numbers of bottles/deposits per good, I can analyze potential effect differences stemming from the number of deposits. In the model, the number of bottles does not affect households' decisions because it was assumed that all relevant marginal benefits and costs were constant. However, we may still see a differentiated effect on the 24 pack as compared to the 12 pack given differences in the price to deposit ratios. There may also be non-modeled behavioral effects, e.g., the salience of a \$2.40 versus \$1.20 redemption potential. The bundle was also designed to tease out the differing tax difference effect between sodas and beer. As previously discussed, Michigan exempts soda from taxation while Indiana and Ohio do not. Beer is taxed in all three states.[23]

Table 1.4 provides summary statistics on pre-tax and pre-deposit retail prices across the

[23] The price of an 11 ounce bag of Doritos was originally collected for usage as a control good. It is dropped due to a lack of variation in prices (suggested retail price printed on each bag), and a near perfect correlation between being on sale and retail type. The control therefore generated no additional predictive power in describing the impact of distance on prices. Gum was also considered as a potential control but Indiana state laws prohibit the sale of gum in liquor stores.

three states. Note that prices in the empirical section are denoted in dollars, not cents. Most importantly, we can see that there is variation in the prices. This variation is fairly consistent across the three states. The lone exception is the case of 24 packs in Ohio where the spread is much smaller.

Table 1.4: Summary of Pre-Tax and Pre-Deposit Retail Prices ($)

Prices ($)	2 Liter Coke	12 Pack Coke	6 Pack Bud	24 Pack BudLite
Indiana				
Mean	1.66	4.66	6.12	16.91
S. Deviation	0.36	0.91	0.51	2.02
Range	[1, 2.49]	[2.50, 5.99]	[5.47, 7.50]	[14.99, 24.91]
Michigan				
Mean	1.78	4.81	6.46	17.53
S. Deviation	0.37	0.94	0.58	1.74
Range	[1, 2.49]	[3.33, 6.99]	[5.19, 8.99]	[14.99, 22.99]
Ohio				
Mean	1.73	4.57	6.23	18.07
S. Deviation	0.30	0.98	0.48	0.58
Range	[1.18, 2.29]	[3, 6.98]	[5.19, 9.10]	[14.99, 19.99]
Avg. Per Bottle Price	1.72	0.39	1.05	0.73

1.4.2 Empirical Design

To estimate the price-distance trends, I use the specification given by (1.4.1). Equation (1.4.1) is run separately for each state sub-sample to estimate the price-distance trends in the three states. With four goods and three states, this gives us a total of 12 regressions. This specification looks specifically at the differential price-distance trend between supermarkets and non-supermarkets for reasons to be discussed below. The coefficients of interest are therefore β_1 and β_3. β_1 is the price change in cents for a 1% increase in driving time while β_3 is the differential price change for supermarkets. The effect of distance on prices is therefore given by β_1 and $\beta_1 + \beta_3$ for non-supermarkets and supermarkets, respectively.

$$\begin{aligned} p_i = \quad & \beta_0 + \beta_1 lnminutes_i + \beta_2 spmkt_i + \beta_3 lnminutes_i * spmkt_i + \beta_4 \mathbf{X}_i \\ & \qquad + \beta_5 sale_i + \beta_6 month_i \end{aligned} \tag{1.4.1}$$

p_i: pre-tax, pre-deposit retail price	$\mathbf{X}_i$: vector of controls
$lnminutes_i$: log driving time (min)	$sale_i$: on sale dummy
$spmkt_i$: supermarket dummy	$month_i$: categorical month dummy

In (1.4.1), the dependent variable is the pre-tax, pre-deposit retail price of the bottled good at given retail location i. The central independent variable, $lnminutes_i$, is the measure of driving time from retailer i to the Michigan border in minutes. Recall that this was calculated using Google Maps. As opposed to driving distance, driving time takes into account differences in the rate of travel as well as abstracting away from potential geographic differences in fuel efficiency and/or costs.[24] Therefore, "distance" will always refer to the driving time in minutes. Due to the distribution and range of driving times in the sample,

[24]The mean minutes-mile difference is fairly small at 1.78 minutes/miles.

I use log minutes in the specifications. This is also appropriate given the asymptotic, non-linear behavior in the model.[25]

I categorize retailers into convenience stores (includes gas stations), grocery stores, liquor stores, pharmacies, and supermarkets. While liquor stores and pharmacies are fairly straightforward to categorize the other three have subtler, more subjective distinctions. I define convenience stores, e.g., 7-Elevens and Speedways, to be those establishments that sold a variety of pre-packaged foods but no fresh produce. Those that did sell produce were classified as grocery stores. To be classified as a supermarket, retailers need to satisfy two requirements. First, they need to sell an expanded selection of foods, e.g., bakery, deli, seafood, etc. Second, the retailer must have at least five chain locations. In these three states, the main and biggest supermarket chains are Meijer, Kroger, and (Super) Wal-Mart. Separating retailers into categories and interacting each with distance allows for systemic differences in price-distance trends. Using the full range of retail categories, however, greatly reduces the power. I therefore use the $spmkt_i$ dummy which separates retailers into supermarkets and non-supermarkets. Supermarket chains tend to follow centralized pricing strategies as evidenced by the homogeneity of advertisements over large geographic areas. We therefore expect them to be less reactive over distance. Separating out this group better allows us to identify a pricing pattern amongst those retailers most likely to have a price-distance trend.

$\mathbf{X_i}$ is a vector of logged demographic and economic controls including population, median household income, median home value, population density, and retail establishment density (proxy for the level of economic activity and firm competition). Retail establishment, in this case, refers to the broader Census definition of retailers and is not specific to vendors of bottled goods. All controls are defined at the standardized zip code level. These controls

[25]The price-distance trends are linear in the simulation results because there are only two cities on either side. With more cities per state, I would expect prices to asymptote towards $\tilde{p}$ on either side.

are meant to account for price patterns caused by concurrent changes in demand and supply conditions. For example, if towns closer to the Michigan border are typically smaller, have fewer stores, and exhibit greater monopoly power, then not accounting for this fact would result in negatively biased estimates of the distance coefficient. For Ohio, I also include the combined local and state sales tax rates.

The $sale_i$ term is a dummy indicating whether a given price observation was marked as being on sale. Note that being *marked* as on sale does not necessarily mean that a given price is in fact on sale, i.e., selling below trend. If a retailer routinely offers the same sale price then it is effectively not on sale.[26] Absent pricing data over time, I am unable to determine the true nature of a given price. This is an issue if, for example, the proportion of true sales to recurring sales is systemically smaller at the border. Regressing with the $sale_i$ term would underestimate the price level at the border and negatively bias the price-distance trend and vice versa if the opposite were true. Therefore, I include results from regressions with and without the $sales_i$ term. The "true" estimate is likely to be somewhere in between. Lastly, the $month_i$ term controls for seasonal factors.[27]

$$\begin{aligned} p_i = \; & \gamma_0 + \gamma_1 lnminutes_i + \gamma_2 spmkt_i + \gamma_3 mich_i + \gamma_4 lnminutes_i * spmkt_i \\ & + \gamma_5 lnminutes_i * mich + \gamma_6 spmkt * mich + \gamma_7 lnminutes_i \\ & * spmkt_i * mich_i + \gamma_8 \mathbf{X}_i + \gamma_9 sale_i + \gamma_{10} month_i \end{aligned} \tag{1.4.2}$$

To estimate the border gap, I use the specification given by (1.4.2). Unlike (1.4.1),

[26] A related example is the strategy of the home goods retailer Bed, Bath, and Beyond which consistently sends out 20% off coupons. A majority of consumers therefore pay 80% of the listed retail prices even though they are technically paying a discounted amount.

[27] A differential test of weekend versus weekday was statistically insignificant.

(1.4.2) is run separately for the Indiana-Michigan and Ohio-Michigan sub-samples. In these two cases, the Michigan sample is divided into Western (Indiana side) and Eastern (Ohio side) Michigan. The Indiana-Ohio border is located almost exactly halfway along Michigan's southern border. Drawing a line straight north from the border provides a cutoff to delineate the two sub-samples. With four goods and two borders, this gives us eight regressions. In the same vein as (1.4.1), (1.4.2) interacts $lnminutes_i$ and $spmkt_i$. It also interacts $mich_i$, a Michigan dummy. This triple interaction, along with the corresponding dual interactions and main effects, allows us to concisely capture the magnitude of the border gap. The coefficients of interest in this case are γ_3 and $\gamma_3 + \gamma_6$ which represent the estimated non-supermarket and supermarket border gaps for Michigan and Indiana/Ohio.[28]

1.4.3 Price-Distance Trends

Tables 1.11 - 1.14 present regression results from (1.4.1), which isolated for a differential effect between supermarkets and non-supermarkets. The non-supermarket price-distance trend is given by the coefficient on $lnminutes_i$, i.e., β_1. For supermarkets, the price-distance trend is given by the sum of β_1 and β_3, the coefficient on the interaction term $lnminutes_i * spmkt_i$. Tables 1.5 and 1.6 condense and highlight the pertinent estimates from each regression.

Let us first consider the non-supermarket prices. Aside from six packs of Budweiser in Ohio, the general price-distance trends are negative in Michigan and positive in Indiana and Ohio. This implies that prices are decreasing with distance from the border in Michigan, but increasing with distance in Indiana and Ohio. These price-distance trends suggest that we are in the intermediate range of household mobility - see Figure 1.3.2.

[28] We can also compare price-distance trend estimates between the two specifications. Estimates are fairly consistent.

Table 1.5: Price-Distance Coefficients for Non-Supermarkets

Non-Spmkts	Michigan	Indiana	Ohio
No Sales			
2L Coke	-0.058*	0.059**	0.044*
12 Coke	-0.087	0.056	0.219**
6 Bud	-0.025	0.104**	-0.093*
24 BudLite	-0.280**	0.049	0.138***
With Sales			
2L Coke	-0.048	0.068***	0.058***
12 Coke	-0.109	0.030	0.154**
6 Bud	-0.025	0.107**	-0.093*
24 BudLite	-0.241*	0.043	0.152***

In comparing the coefficients from the sales and no-sales regressions, we see that the inclusion of the $sale_i$ term does not consistently shift the price-distance trend up or down.[29] Comparing effects across the different products, we see that going from a good with fewer deposits to a good with more deposits (within the two types) decreases the coefficients in Michigan and Indiana but increases the coefficients in Ohio for the non-supermarket prices. The simulation results suggested that an increase in the deposit would increase and decrease the price-trend in Michigan and Indiana, respectively. In this case, only Indiana follows the predicted pattern. While the supermarket coefficients seem to fit the model's predictions

[29]Table 1.10 shows the estimated marginal effects from a probit regression which looks at the probability of seeing a sale as a function of distance. Sales tend to be less prevalent closer to the border for two liters of Coke and 24 packs of Bud Lite, and more prevalent for 12 packs of Coke. There does appear to be a relationship between the two but R^2 values between 0.20-0.30 suggest that collinearity is not a significant issue.

more so than the non-supermarket prices, overall it seems that this prediction does not hold. However, it is unclear whether or not this comparison should be made since only the per unit price to deposit ratios are increasing, and not the actual deposit. Likewise, comparisons between sodas and beers are tenuous due to differences in the number of bottle deposits.[30]

Table 1.6: Price-Distance Coefficients for Supermarkets

Spmkts	Michigan	Indiana	Ohio
No Sales			
2L Coke	-0.064**	0.034	0.030
12 Coke	0.184**	-0.424***	0.014
6 Bud	0.155	0.025	-0.154
24 BudLite	0.080	-0.320	0.088
With Sales			
2L Coke	0.033	0.014	0.032
12 Coke	0.314***	-0.235**	0.086
6 Bud	0.155	0.026	-0.154
24 BudLite	0.364	-0.270	0.103

If we compare effects between supermarkets and non-supermarkets, we confirm that supermarket prices tend to react less with distance from the border than non-supermarkets. This is true for all goods except 12 packs of Coke. In Ohio, all estimates are insignificant. If we include sales in the regressions, the estimates are insignificant in all but two cases, again for 12 packs of Coke. As discussed previously, the non-responsiveness of supermarket

[30] A better comparison may have been possible if I had collected data on soda and beer products with equal deposits, e.g., a six pack of Coke versus a six pack of Budweiser. However, single bottles of beer, six packs of soda, and 24 packs of soda are not commonly found.

prices over distance is likely due to the larger, retail chain nature of these stores. The popularity of these four products guarantees that they are almost always advertised and/or on sale. Because the same weekly advertisements go out for all stores within a given region, this hinders the ability of individual stores to tailor pricing decisions at individual locations. It may also limit such behavior to the pricing of less popular and/or regional items. Non-supermarkets, in particular non-chain retailers, may be more decentralized and better able to adjust to location specific factors.[31] This is very true in the case of the big three supermarket chains, Kroger, Meijer, and Wal-Mart, where differences in advertised prices only exist across different (large) regions. It is also possible that supermarkets are more competitive than non-supermarkets. If this were true, then we would also expect to see less of a price-distance reaction.

Since the Michigan sample includes retailers near both borders, it is possible that the Michigan regressions are aggregating two statistically different price-distance responses. If Indiana and Ohio are the same, then this is not an issue. However, if Indiana and Ohio, or East and West Michigan are different, then this aggregation could be clouding the results. To test this possibility, Table 1.7 summarizes estimates from separate regressions on the two Michigan sub-samples. Specifically, I divide up the Michigan sample into those in East versus West Michigan. The north-south, Indiana-Ohio break coincides with the county boundaries in Michigan. Comparing the two sub-samples, we see that differences between the coefficients are only statistically significant for two liter of Cokes (non-supermarket with sales) and 24 packs of Bud Lite (all). In comparison with the estimates from Tables 1.5 and 1.6, the largest change occurs in the non-supermarket Bud Lite regressions. Previously, the combined Michigan sample generated an estimate of -0.280 and -0.241 for the no sale

[31]Using a $chain_i$ dummy instead of $spmkt_i$ produces less statistically significant estimates.

and sale regressions, respectively. With the split sample, West Michigan has a negative estimate while East Michigan has a positive estimate. For the supermarket prices, there are some differences in sign although these differences only occur where there is no statistical significance.

Table 1.7: East Versus West Michigan

	Non-Supermarkets		**Supermarkets**	
No Sales	West Mich	East Mich	West Mich	East Mich
2L Coke	-0.106**	0.013	-0.087	-0.050
12 Coke	-0.114	-0.066	0.273*	0.134
6 Bud	-0.039	-0.140	0.101	-0.042
24 BudLite	-0.584**	0.684	0.079	-0.063
With Sales				
2L Coke	-0.075*	0.005	0.103	-0.007
12 Coke	-0.205	-0.038	0.365***	0.295***
6 Bud	-0.039	-0.140	0.101	-0.042
24 BudLite	-0.542**	0.815*	0.416*	0.332

There are two plausible explanations for this disparity between the non-supermarket price-distance trends. First, there could be a significant difference in the characteristics of households between the Indiana and Ohio sides. For example, households in the west could be less mobile than households in the east, which the simulated model shows could indeed flip the signs. There could also be differences in household preferences/demand due to the economic profile of border regions in the two sides. As previously discussed, West Michigan is fairly rural near the border with some tourism based towns surrounding Lake

Michigan. The negative coefficient in Indiana could be driven by a higher willingness to pay in West Michigan due to the presence of wealthier non-resident tourists whose (transient) incomes would not be captured by the controls. Second, these results could also be caused by differences in alcohol laws between Indiana and Ohio. Specifically, Indiana bans the sale of carry-out beer on Sundays from all retail establishments. Given the close proximity of the South Bend area to the Indiana-Michigan border, this Sunday ban may push greater demand into Michigan. This additional wedge would support the decreasing price-distance trend because Michigan stores close to the border could exploit their Sunday monopoly. Ohio also has alcohol laws although these tend to cover hard alcohol sales rather than beer sales. Additionally, these laws are set at the local level where they are primarily found in rural towns further from the border. Due to the fact that the two liter Cokes also exhibit this pattern, albeit at a much weaker level, it is likely that the true explanation involves a combination of these two explanations.

1.4.4 Border Gaps

Given the estimated price-distance trends and the results from the simulated model, we would expect to see a negative border gap, i.e., Indiana and Ohio's border prices being lower than Michigan's border prices. We can verify this by looking at Tables 1.17 and 1.18 which present estimation results from specification (1.4.2). Recall that these regressions include interactions between $lnminutes_i$, $spmkt_i$, and a Michigan dummy, $mich_i$. Running this specification over the combined Indiana and Michigan (West/Indiana side), and Ohio and Michigan (East/Ohio side) samples gives us a straightforward way to quantify the border gaps. Our estimates of interest are γ_3, the coefficient on the Michigan dummy, and γ_6, the coefficient on the interaction between the Michigan and supermarket dummies. γ_3 is the non-supermarket, border price difference between Michigan and Indiana/Ohio while γ_6 is the additional supermarket price difference. The border gaps for non-supermarkets and supermarkets are therefore given by $-\gamma_3$ and $-(\gamma_3 + \gamma_6)$, respectively. Technically, these are the border price differentials when $lnminutes_i$ is equal to zero, i.e., when the driving time is equal to one minute.[32]

Table 1.8: Border Gap Coefficients

	Non-Supermarkets		**Supermarkets**	
No Sales	Ind-Mich	Ohio-Mich	Ind-Mich	Ohio-Mich
2L Coke	-0.447***	-0.184	-0.659**	-0.255
12 Coke	-0.580	-0.341	1.677**	0.858*
6 Bud	-0.230	-0.762	0.290	0.384
24 BudLite	-1.667*	1.038	0.854	0.746
With Sales				
2L Coke	-0.374**	-0.133	0.081	-0.199
12 Coke	-0.672	-0.409	1.718**	0.920**
6 Bud	-0.247	-0.762	0.280	0.384
24 BudLite	-1.569	1.229*	1.987	1.111**

Table 1.8 summarizes the border gap estimates from these regressions as well as regressions without the $sale_i$ term. Notice that in all cases, except for 24 packs of Bud Lite at the Ohio-Michigan border, the border gap is negative, i.e., the Michigan prices are higher at the border. The -0.374 border gap estimate for two liters implies that Indiana non-supermarkets

[32]Depending on the relative steepness of the price-distance trends, these estimates could be greater or less than the true border gaps because we are measuring the difference at one minute of driving time. However, these estimates are not likely to be far from the true border gap.

on the border are on average 37 cents cheaper than their Michigan counterparts. Similar to the pattern found in the price-distance trend regressions, these results are consistent with household mobility being in the intermediate range - see Figure 1.3.2. However, only the estimates for the two liter Cokes and the 24 pack of Bud Lite (sales only) are significant.

In regards to the supermarket border gaps, notice that most, with the exception of two liter Cokes, are now positive, i.e., border prices in Indiana and Ohio are greater than border prices in Michigan. From the price-distance trend regressions, we saw that supermarkets are typically less responsive within each state due to the more centralized, chain pricing strategies. However, this is less true across states as the big three offer different advertisements and/or sales between states and even larger regions such as East and West Michigan. Additionally, there exists a number of more regional supermarket chains, e.g., Martin's Supermarkets in Indiana, Chief and Giant Eagle Supermarkets in Ohio, and Harding's Supermarket in West Michigan, that service the border areas in these three states. Like the big three, their prices are constant from store to store. Due to their slightly higher prices (relative to the big three) and their regional nature near the border, these smaller supermarkets make it look as if supermarkets as a whole are more expensive near the border. Thus, we see a number of these positive border gaps in the supermarket category. If we redefine the $spmkt_i$ term to exclude these smaller supermarkets, then these coefficients decrease in both magnitude and significance. For example, the \$1.68 border gap for the 12 packs of Coke on the Indiana-Michigan border is no longer significant. On the non-supermarket side, the addition of these stores shrinks the border gap from -\$0.45 to -\$0.16 with both being significant.

The exceptions to these general trends are the non-supermarkets prices of 24 packs and the supermarket prices of two liter Cokes. However, it makes intuitive sense that the 24 pack would have a positive border gap due to its relatively high number of deposits, low price-to-deposit ratio, and small tax differential. We would expect the bottle bill wedge to be strongest for this good. Likewise for two liters of Cokes in supermarkets, the good with the lowest deposit, highest price-to-deposit ratio, and highest tax differential has a negative gap.

While the border gaps in the non-supermarket regressions tend to match with the model predictions (given the price-distance trends), we only find limited significance. This could be an issue related to the sample size. It could also be an issue with noise and/or mis-measurement in the data, e.g., the indeterminate nature of the $sale_i$ term. Alternatively, it is possible that $-\gamma_3$ and $-(\gamma_3 + \gamma_6)$ are not our estimates of interest. For a consumer interested in cross-border shopping, the price - not the price differential conditional on being a supermarket or non-supermarket - is the parameter of interest. The model assumed that retailers' locations were exogenously given. This did not matter in terms of the simulation results as all firms were assumed identical. In reality, however, retailers are not identical and the location choice of a given retailer/retail type is endogenous. This suggests that it may be appropriate to *not* control for retail type. The same argument could be made for excluding the $sale_i$ term if we believe that households can time their purchases to coincide with these sales.

Similar to the argument for splitting Michigan into East and West Michigan, specification (1.4.2) also looks to identify estimates of the border gap across the entirety of the Indiana-Michigan and Ohio-Michigan borders. It is plausible to believe that such a discontinuity may be more readily apparent if we examine specific areas with greater economic activity, traffic, and retailers. Specification (1.4.3) therefore presents an alternative strategy to search for price discontinuities at the two specific borders near Toledo, Ohio and South Bend, Indiana.

$$\begin{aligned} p_i = \quad & \delta_0 + \delta_1 close_i + \delta_2 mich_i + \delta_3 close_i * mich_i + \delta_4 lnminutes_i * (1 - close_i) \\ & + \delta_5 lnminutes_i * mich * (1 - close_i) + \bar{\delta_6}\mathbf{X}_i + \delta_7 sale_i \end{aligned} \tag{1.4.3}$$

For the Toledo border, the sample is limited to observations from the counties of Lucas (OH) and Monroe (MI). For the South Bend border, the sample is limited to observations from the counties of Elkhart (IN), La Porte (IN), St. Joseph (IN), Berrien (MI), and Cass (MI). The variable $close_i$ is a dummy which equals one if the retailer is five minutes or closer to the border and zero otherwise. This specification differs from the previous specifications in that it aggregates prices within this five minutes range, i.e., it assumes that prices do not react substantially with respect to distance in this narrow time range. Only for locations further away do we introduce the $lnminutes_i$ term. Additionally, I no longer distinguish between supermarkets and non-supermarkets. Due to the sample size, only controls for logged median household income, population density, and retail establishment density are included. The border gaps are therefore given by $-(\delta_2 + \delta_3)$.

Table 1.9: Summary Table of Border Gap Coefficients - Combined

All Retailers	South Bend, IN	Toledo, OH
With Sales		
2L Coke	-0.333***	-0.171
12 Coke	1.081***	-0.766*
6 Bud	-0.851**	-0.059
24 BudLite	-1.213	1.168***

Tables 1.19 and 1.20 present the regression results for these two border cases with border

gap estimates shown in the bottom two rows. Table 1.9 summarizes these estimates. In comparison to the border gaps from the previous specification, we see that the estimates are more significant. This does suggest that the impact of the bottle bill and tax wedges are likely to be more focused at these concentrated areas. Combining supermarkets and non-supermarkets also means that the estimated border gaps are sandwiched in between the border gaps from Table 1.8. They are typically closer to the non-supermarket estimates due to the greater number of non-supermarket observations.

1.4.5 Sensitivity Analysis

To test the sensitivity of the price-trend results from the different specifications, I also ran the regressions with city level controls instead of zip code level controls, which generated negligible differences.[33] This also occurred when using logged driving distances instead of logged driving times. Because the number of clusters is less than 50 in the single state regressions, I also re-calculated the standard errors using wild bootstraps as suggested by Cameron et al. [2008].[34] This generated mixed effects as errors decreased for some coefficients while they increased in others.

1.5 Conclusion

This paper analyzed the capitalization of a bottle bill and tax wedge in the patterns of retail bottle prices near the borders of Indiana, Michigan, and Ohio using both a theoretical model and empirical analysis of an original dataset. By quantifying the price-distance trends and

[33] I did not have retail establishment data at the city level but did include a control for the number of supermarkets and club membership stores within city boundaries.

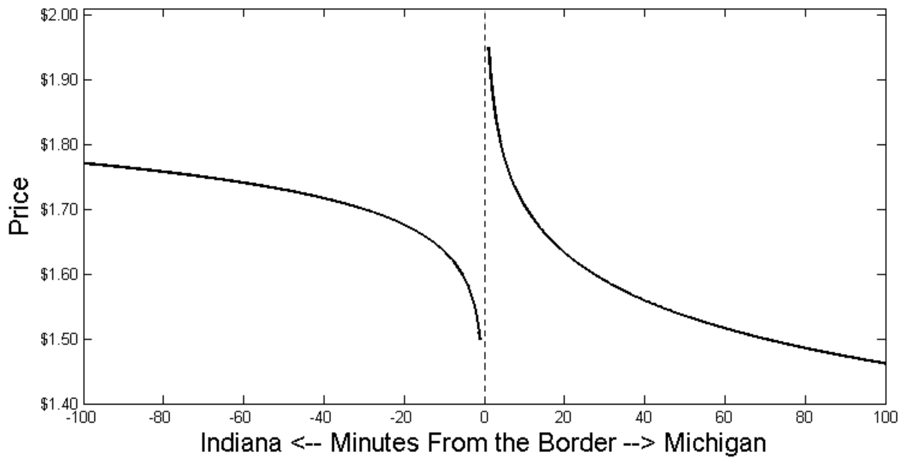
[34] Typically, clustering improves standard errors but may be detrimental if there are fewer than 50 clusters.

border gaps in these three states, I indirectly identified evidence of cross-border shopping, fraudulent redemptions, and use tax evasion by households.

From a simulated model of imperfect intercity and interstate price competition, I generated predictions on the patterns of equilibrium prices and the relative strengths between the two wedges. The findings from the model suggest that the degree of household mobility is the main determinant of both the price-distance trends and border gaps. High levels of mobility result in lower equilibrium prices, particularly at the borders where competition is the most intense. At these low prices, a percentage tax difference is also weaker in strength relative to the deposit. Thus, price-distance trends are positive while the border gap is negative. As travel costs increase, these results flip.

Using an original dataset of retail prices, I then tested the implications of the model by empirically quantifying these border gaps and price-distance trends for four bottled goods. In general, the empirical results were consistent across the different specifications. Near the Indiana-Michigan border, the non-supermarket price-distance trends for two liters of Coke were estimated to be 0.059 and -0.106, respectively. Given an average price of $1.72 for the Cokes, these estimates imply that prices at non-supermarkets located 30 minutes from the border are $0.20 more expensive and $0.36 cheaper for Indiana and Michigan, respectively. Overall, non-supermarket prices in Indiana and Ohio tend to increase with distance from the border whereas prices in Michigan tend to decrease with distance from the border. Supermarket price-trends are less significant and smaller in magnitude with the exception of 12 packs of Coke. This result coincides with the homogeneity of advertised prices across their chain stores.

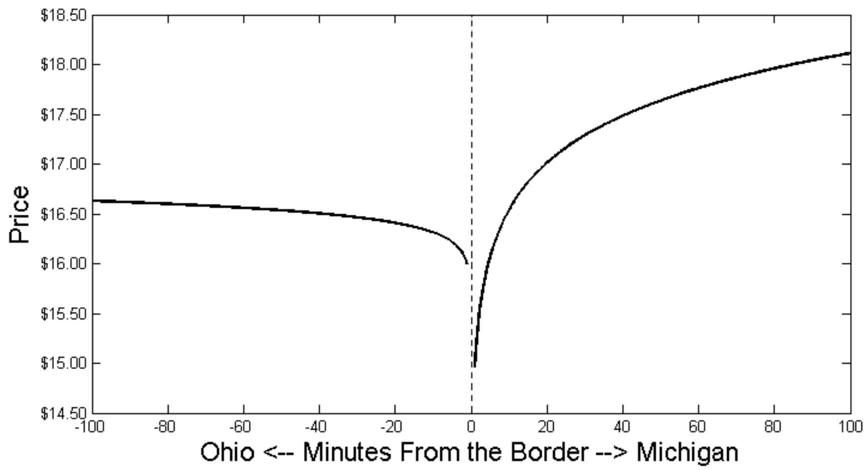
Figure 1.5.1: Non-Supermarket 2L Coke Price Trends and Gap (IN-MI Border)



Note: Price levels are normalized at $1.50 for the Indiana border. Covariates are fixed at all distances.

Given the estimated price-distance trends, the estimated border gaps matched up well with the predictions of the model. In general, Michigan's border prices were found to be higher (lower) than those of non-supermarkets (supermarkets) in Indiana and Ohio. This suggests that the effect of the bottle bill wedge is weaker relative to the tax wedge. Returning to the two liter Coke example, I estimated a -$0.45 border gap. Given that the 7% sales tax difference between the two states generates an average tax hit of approximately $0.11, whereas purchasing in Indiana saves only $0.10 via the deposit, it appears that retail prices are reacting more to the tax wedge. Likewise, I estimated a border gap of -$1.67 for 24 packs of Bud Lite. This produces a picture similar to Figure 1.5.1. However, the same border gap was estimated to be $1.04 on the Ohio-Michigan border. In this case, the positive border gap indicates that the bottle bill wedge, particularly for high deposit and/or low price-to-deposit ratio goods, can dominate the tax wedge. It also highlights the presence of additional factors, e.g., Indiana's Sunday alcohol ban.

Figure 1.5.2: Non-Supermarket 24 BudLite Price Trends and Gap (OH-MI Border)



Note: Price levels are normalized at $16.00 for the Ohio border. Covariates are fixed at all distances.

Overall, the results indirectly suggest the presence of cross-border shopping, fraudulent redemptions, and use tax evasion. Given the estimated border gaps and price-distance trends, prices reflect a greater impact from the tax differential and use tax evasion. Note, however, that this does not indicate the absence of cross-border shopping and/or fraudulent redemptions, only the relative nature and capitalization of such behavior in the retail prices. Given the lack of direct data, this indirect empirical approach is a feasible strategy for identifying and quantifying the relative strengths of these illegal activities. We can also apply this methodology to other state borders such as Maine which has had documented cases of organized fraudulent redemptions New Hampshire. New York has an even greater border with the non-bottle bill state of Pennsylvania. Additionally, New York City has a smaller but dense border with New Jersey, both in terms of population and retail establishments. As such, these other cases present opportunities for further research and study of this important policy issue.

Table 1.10: Marginal Effects of a Probit Regression Examining Sales and Distance

Sale	(1) 2LCoke	(2) 12Coke	(3) 24BudLite
Spmkt	0.38369***	0.48068***	0.41498***
	(0.090)	(0.032)	(0.038)
Michigan	-0.70121***	0.14693	-0.72386***
	(0.092)	(0.164)	(0.044)
Ohio	-0.25167***	0.07943***	-0.23290***
	(0.026)	(0.029)	(0.050)
Log Minutes	0.00244	0.06594*	0.01499
	(0.006)	(0.036)	(0.011)
Mich*Log Minutes	0.11622***	-0.04601*	0.09776***
	(0.021)	(0.026)	(0.006)
Ohio*Log Minutes	0.01000	-0.05059***	0.02089**
	(0.009)	(0.007)	(0.009)
Sales Tax (Per.Points)	-0.66015***	-0.02542	-0.46935***
	(0.063)	(0.064)	(0.088)
Log Total Pop	-0.00710	0.01110	0.03412
	(0.038)	(0.013)	(0.037)
Log Med. HH Income	0.17112	-0.11658*	0.15494***
	(0.194)	(0.063)	(0.032)
Log Med. Home Value	0.02369	0.14430*	-0.09538**
	(0.136)	(0.081)	(0.037)
Log Density	-0.00661	-0.04427***	-0.02840
	(0.044)	(0.010)	(0.036)
Log Retail Density	0.01570	0.02987	0.05154***
	(0.050)	(0.029)	(0.013)
Observations	346	346	337

Robust standard errors are in parentheses and clustered by state. Baseline is an Indiana non-supermarket in February. Month coefficients omitted.

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

Table 1.11: 2 Liter Coke Regressions with Spmkt Dummy

2L Coke	(1) Michigan	(2) Michigan	(3) Indiana	(4) Indiana	(5) Ohio	(6) Ohio
Spmkt	-0.34565**	-0.38208	-0.39061***	-0.02015	-0.40317***	-0.28504***
	(0.125)	(0.231)	(0.089)	(0.117)	(0.043)	(0.054)
Log Minutes	-0.05844*	-0.04836	0.05911**	0.06831***	0.04425*	0.05767***
	(0.031)	(0.030)	(0.025)	(0.018)	(0.024)	(0.019)
Log Minutes*Spmkt	-0.00557	0.08127	-0.02494	-0.05448	-0.01466	-0.02569
	(0.042)	(0.082)	(0.042)	(0.036)	(0.029)	(0.023)
Log Total Pop	-0.07036*	-0.06956***	0.01892	-0.01580	-0.05304	-0.04477
	(0.039)	(0.024)	(0.039)	(0.036)	(0.040)	(0.034)
Log Med. HH Income	-0.12940	-0.08872	0.14758	0.13608	-0.03495	0.07999
	(0.113)	(0.111)	(0.419)	(0.290)	(0.229)	(0.223)
Log Med. Home Value	-0.09977	-0.06971	0.11803	0.10348	-0.02852	-0.09209
	(0.068)	(0.081)	(0.355)	(0.248)	(0.214)	(0.217)
Log Density	0.09541**	0.08258***	0.01579	0.04279*	0.04021	0.01719
	(0.037)	(0.023)	(0.029)	(0.022)	(0.032)	(0.026)
Log Retail Density	-0.07749**	-0.05858***	0.00507	-0.03273*	-0.03274	-0.00310
	(0.032)	(0.018)	(0.028)	(0.017)	(0.036)	(0.030)
Sales Tax (Per.Points)					-0.01229	-0.09185
					(0.152)	(0.122)
Sale		-0.48449***		-0.44763***		-0.25989***
		(0.064)		(0.085)		(0.069)
Constant	4.77671***	4.05042***	-1.60050	-1.07650	2.89297*	2.85725**
	(0.712)	(0.686)	(1.281)	(1.323)	(1.418)	(1.164)
Observations	108	108	99	99	86	86
R-squared	0.476	0.679	0.362	0.549	0.476	0.560
Spmkt Trend	-0.064	0.033	0.034	0.014	0.030	0.032
P-Value	0.023	0.590	0.406	0.682	0.280	0.175

Robust standard errors are in parentheses and clustered by county. Baseline is a non-supermarket in February. Month coefficients omitted.

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

Table 1.12: 12 Pack Coke Regressions with Spmkt Dummy

12Coke	(1) Michigan	(2) Michigan	(3) Indiana	(4) Indiana	(5) Ohio	(6) Ohio
Spmkt	-1.97936***	-2.12840***	0.91837	0.70598*	-0.66540***	-0.65609***
	(0.311)	(0.259)	(0.550)	(0.405)	(0.158)	(0.143)
Log Minutes	-0.08714	-0.10948	0.05646	0.03001	0.21918**	0.15444**
	(0.130)	(0.121)	(0.071)	(0.068)	(0.084)	(0.066)
Log Minutes*Spmkt	0.27113**	0.42377***	-0.48025***	-0.26531*	-0.20507**	-0.06820
	(0.126)	(0.111)	(0.168)	(0.129)	(0.081)	(0.068)
Log Total Pop	-0.11056	-0.07256	-0.09103	-0.08810	-0.09005	-0.07285
	(0.109)	(0.093)	(0.121)	(0.109)	(0.115)	(0.116)
Log Med. HH Income	0.24881	0.23590	0.89582	0.72611	0.10334	0.29664
	(0.533)	(0.480)	(0.987)	(0.571)	(0.433)	(0.672)
Log Med. Home Value	-0.95707	-0.77653	-0.73118	-0.75517*	0.24335	-0.02193
	(0.571)	(0.491)	(0.549)	(0.375)	(0.433)	(0.623)
Log Density	0.18435	0.11449	0.14976	0.11508	0.13278	0.06421
	(0.122)	(0.101)	(0.103)	(0.086)	(0.077)	(0.090)
Log Retail Density	-0.20603*	-0.12718	-0.10128	-0.11613*	-0.04987	0.01373
	(0.108)	(0.097)	(0.070)	(0.056)	(0.042)	(0.054)
Sales Tax (Per.Points)					0.20612	0.26945
					(0.515)	(0.398)
Sale		-0.79203***		-0.99532***		-0.90346***
		(0.135)		(0.126)		(0.154)
Constant	14.35206***	12.55008***	3.94864	6.33693	-1.01403	-0.15237
	(3.045)	(2.581)	(6.111)	(5.058)	(4.500)	(3.827)
Observations	91	91	92	92	81	81
R-squared	0.627	0.726	0.437	0.593	0.432	0.543
Spmkt Trend	0.184	0.314	-0.424	-0.235	0.014	0.086
P-Value	0.021	0.000	0.004	0.021	0.816	0.176

Robust standard errors are in parentheses and clustered by county. Baseline is a non-supermarket in February. Month coefficients omitted.

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

Table 1.13: 6 Pack Budweiser Regressions with Spmkt Dummy

6Bud	(1) Michigan	(2) Michigan	(3) Indiana	(4) Indiana	(5) Ohio	(6) Ohio
Spmkt	-1.05713***	-1.05713***	-0.37015	-0.35133	-0.21040	-0.21040
	(0.290)	(0.290)	(0.327)	(0.334)	(0.187)	(0.187)
Log Minutes	-0.02489	-0.02489	0.10380**	0.10681**	-0.09322*	-0.09322*
	(0.041)	(0.041)	(0.041)	(0.039)	(0.048)	(0.048)
Log Minutes*Spmkt	0.17958*	0.17958*	-0.07878	-0.08092	-0.06056	-0.06056
	(0.100)	(0.100)	(0.121)	(0.125)	(0.139)	(0.139)
Log Total Pop	0.06805	0.06805	-0.26772*	-0.27847*	0.09657	0.09657
	(0.068)	(0.068)	(0.137)	(0.136)	(0.110)	(0.110)
Log Med. HH Income	0.45107**	0.45107**	0.41840	0.39923	0.45894	0.45894
	(0.166)	(0.166)	(0.489)	(0.490)	(0.472)	(0.472)
Log Med. Home Value	-0.43864**	-0.43864**	-1.05189	-1.04681	-0.36032	-0.36032
	(0.196)	(0.196)	(0.699)	(0.701)	(0.332)	(0.332)
Log Density	-0.03387	-0.03387	0.30807***	0.31798***	-0.07347	-0.07347
	(0.060)	(0.060)	(0.103)	(0.103)	(0.119)	(0.119)
Log Retail Density	0.08487*	0.08487*	-0.28235**	-0.28140**	-0.01548	-0.01548
	(0.047)	(0.047)	(0.109)	(0.110)	(0.045)	(0.045)
Sales Tax (Per.Points)					0.31028	0.31028
					(0.259)	(0.259)
Sale				0.17464		
				(0.428)		
Constant	6.40759***	6.40759***	14.99149**	15.16099**	3.35610*	3.35610*
	(2.038)	(2.038)	(6.018)	(6.001)	(1.763)	(1.763)
Observations	97	97	49	49	77	77
R-squared	0.248	0.248	0.658	0.662	0.145	0.145
Spmkt Trend	0.155	0.155	0.025	0.026	-0.154	-0.154
P-Value	0.129	0.129	0.858	0.855	0.356	0.356

Robust standard errors are in parentheses and clustered by county. Baseline is a non-supermarket in February. Month coefficients omitted.

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

Table 1.14: 24 Pack BudLite Regressions with Spmkt Dummy

24BudLite	(1) Michigan	(2) Michigan	(3) Indiana	(4) Indiana	(5) Ohio	(6) Ohio
Spmkt	-2.58814**	-2.84078**	-0.50776	0.12163	0.27918**	0.31766**
	(1.019)	(1.086)	(1.003)	(1.090)	(0.127)	(0.126)
Log Minutes	-0.27975**	-0.24056*	0.04907	0.04339	0.13813***	0.15222***
	(0.120)	(0.122)	(0.145)	(0.161)	(0.041)	(0.047)
Log Minutes*Spmkt	0.35944	0.60416**	-0.36945	-0.31343	-0.04974	-0.04950
	(0.249)	(0.280)	(0.232)	(0.187)	(0.066)	(0.068)
Log Total Pop	-0.12035	-0.00726	-0.22868	-0.20632	-0.00954	-0.01532
	(0.198)	(0.206)	(0.556)	(0.472)	(0.115)	(0.118)
Log Med. HH Income	-0.88654	-0.77047	-3.18802**	-2.83778	-0.93720**	-0.88901*
	(0.995)	(0.974)	(1.473)	(1.715)	(0.437)	(0.423)
Log Med. Home Value	-0.59577	-0.49672	2.48451	2.36221	1.03865**	0.99720**
	(0.691)	(0.700)	(1.689)	(1.593)	(0.439)	(0.424)
Log Density	0.19770	0.07780	0.31845	0.29274	0.09194	0.09473
	(0.190)	(0.197)	(0.614)	(0.514)	(0.088)	(0.091)
Log Retail Density	-0.21677*	-0.11560	-0.33474	-0.24861	-0.08589	-0.08075
	(0.122)	(0.121)	(0.649)	(0.479)	(0.061)	(0.059)
Sales Tax (Per.Points)					0.81658***	0.77428***
					(0.219)	(0.203)
Sale		-1.11746***		-1.48836		-0.14211
		(0.247)		(0.869)		(0.126)
Constant	35.58890***	32.71156***	22.81986**	20.51338**	9.40850***	9.61269***
	(6.655)	(6.423)	(8.155)	(8.234)	(2.492)	(2.472)
Observations	105	105	70	70	91	91
R-squared	0.318	0.349	0.332	0.417	0.210	0.215
Spmkt Trend	0.080	0.364	-0.320	-0.270	0.088	0.103
P-Value	0.716	0.151	0.365	0.337	0.133	0.143

Robust standard errors are in parentheses and clustered by county. Baseline is a non-supermarket in February. Month coefficients omitted.

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

Table 1.15: Michigan Split Sample - 2L Coke Regressions

2LCoke	(1) Indiana Side	(2) Indiana Side	(3) Ohio Side	(4) Ohio Side
Spmkt	-0.28406	-0.61169***	-0.29795*	-0.12717
	(0.186)	(0.153)	(0.141)	(0.157)
Log Minutes	-0.10649**	-0.07515*	0.01334	0.00537
	(0.035)	(0.034)	(0.065)	(0.057)
Log Minutes*Spmkt	0.01908	0.17809**	-0.06378	-0.01194
	(0.059)	(0.062)	(0.046)	(0.053)
Log Total Pop	-0.14489*	-0.04357	-0.04434	-0.11890*
	(0.072)	(0.060)	(0.063)	(0.062)
Log Med. HH Income	-0.05818	-0.13554	-0.06892	-0.37069
	(0.196)	(0.228)	(0.154)	(0.243)
Log Med. Home Value	-0.22854	-0.07673	-0.14252	0.12327
	(0.135)	(0.109)	(0.144)	(0.210)
Log Density	0.12226**	0.03391	0.08056	0.13959**
	(0.052)	(0.047)	(0.072)	(0.059)
Log Retail Density	-0.06057	-0.03481	-0.07476	-0.10195**
	(0.035)	(0.019)	(0.067)	(0.045)
Sale		-0.58217***		-0.44369***
		(0.071)		(0.064)
Constant	6.15576***	4.71688**	4.11719***	4.77662***
	(1.730)	(1.450)	(0.848)	(0.913)
Observations	45	45	63	63
R-squared	0.511	0.764	0.536	0.670
Spmkt Trend	-0.087	0.103	-0.050	-0.007
P-Value	0.320	0.216	0.289	0.895

Robust standard errors are in parentheses and clustered by zip. Baseline is a non-supermarket in February. Month coefficients omitted.

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

Table 1.16: Michigan Split Sample - 24 Bud Lite Regressions

24BudLite	(1) Indiana Side	(2) Indiana Side	(3) Ohio Side	(4) Ohio Side
Spmkt	-3.55503***	-3.95152***	0.76640	0.70286
	(0.543)	(0.726)	(0.723)	(0.882)
Log Minutes	-0.58389**	-0.54158**	0.68422	0.81474*
	(0.212)	(0.225)	(0.425)	(0.401)
Log Minutes*Spmkt	0.66277***	0.95741***	-0.74734**	-0.48265
	(0.193)	(0.222)	(0.272)	(0.346)
Log Total Pop	-0.76153	-0.57521	-0.22536	-0.12747
	(0.438)	(0.508)	(0.318)	(0.291)
Log Med. HH Income	-0.18784	0.02830	-0.79474	-0.67554
	(0.892)	(1.034)	(1.381)	(1.184)
Log Med. Home Value	-1.95225***	-1.75284***	-0.33570	-0.19959
	(0.303)	(0.405)	(1.265)	(1.104)
Log Density	0.51885	0.34065	0.37913	0.26293
	(0.367)	(0.438)	(0.337)	(0.314)
Log Retail Density	-0.31233	-0.18650	-0.32356	-0.21019
	(0.249)	(0.271)	(0.317)	(0.314)
Sale		-1.33671		-1.48121***
		(0.740)		(0.410)
Constant	48.82889***	43.36298***	28.26008***	24.77167***
	(10.039)	(10.146)	(5.844)	(5.059)
Observations	45	45	60	60
R-squared	0.522	0.553	0.332	0.400
Spmkt Trend	0.079	0.416	-0.063	0.332
P-Value	0.743	0.090	0.827	0.293

Robust standard errors are in parentheses and clustered by zip. Baseline is a non-supermarket in February. Month coefficients omitted.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.17: Indiana-Michigan Combined Regression

VARIABLES	(1) 2LCoke-InMi	(2) 12Coke-InMi	(3) 6Bud-InMi	(4) 24BudLite-InMi
Spmkt	0.02194	0.80711*	-0.35894	0.09092
	(0.113)	(0.416)	(0.212)	(1.001)
Log Minutes	0.06660***	0.02793	0.11660**	0.08247
	(0.020)	(0.077)	(0.050)	(0.101)
Michigan	0.37408***	0.67210	0.24698	1.56917
	(0.131)	(0.690)	(0.280)	(1.045)
Mich*Spmkt	-0.45488*	-2.39050***	-0.52733**	-3.55587*
	(0.264)	(0.795)	(0.247)	(1.774)
Spmkt*Log Minutes	-0.06593*	-0.32758**	-0.14025*	-0.28855
	(0.036)	(0.147)	(0.072)	(0.209)
Mich*Log Minutes	-0.11853***	-0.08264	-0.11108	-0.45229**
	(0.037)	(0.293)	(0.086)	(0.200)
Mich*Spmkt*Log Minutes	0.15646*	0.58946*	0.29489***	1.02279**
	(0.086)	(0.302)	(0.091)	(0.476)
Sale	-0.47936***	-0.99242***	0.14170	-1.50599**
	(0.071)	(0.118)	(0.372)	(0.685)
Constant	0.20006	8.02685	7.24679**	28.29351***
	(1.430)	(5.225)	(3.282)	(6.473)
Observations	144	126	93	115
R-squared	0.594	0.578	0.471	0.429
Spmkt Border Gap	0.081	1.718	0.280	1.987
P-Value	0.729	0.003	0.155	0.413

Robust standard errors are in parentheses and clustered by county. Baseline is an Ohio non-supermarket in February. Month coefficients and standard controls omitted.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.18: Ohio-Michigan Combined Regression

VARIABLES	(1) 2LCoke-OhMi	(2) 12Coke-OhMi	(3) 6Bud-OhMi	(4) 24BudLite-OhMi
Spmkt	-0.25251***	-0.58590***	-0.22039	0.56787***
	(0.060)	(0.146)	(0.172)	(0.200)
Log Minutes	0.06645***	0.10991	-0.05377	0.22430***
	(0.021)	(0.070)	(0.048)	(0.060)
Michigan	0.13261	0.40875	0.76233	-1.22901*
	(0.162)	(0.350)	(0.642)	(0.712)
Mich*Spmkt	0.06619	-1.32857***	-1.14665***	0.11825
	(0.083)	(0.248)	(0.414)	(0.498)
Spmkt*Log Minutes	-0.03587	-0.10146	-0.06335	-0.07216
	(0.024)	(0.066)	(0.113)	(0.076)
Mich*Log Minutes	-0.07498**	0.00633	-0.07171	0.43994*
	(0.030)	(0.096)	(0.210)	(0.223)
Mich*Spmkt*Log Minutes	-0.00891	0.33484***	0.25771	-0.56584***
	(0.036)	(0.121)	(0.160)	(0.196)
Sale	-0.30415***	-0.85626***		-0.76052**
	(0.065)	(0.094)		(0.327)
Constant	3.22066***	3.43614	3.32620*	16.34004***
	(1.111)	(3.618)	(1.930)	(5.010)
Observations	149	138	130	151
R-squared	0.588	0.636	0.229	0.366
Spmkt Border Gap	-0.199	0.920	0.384	1.111
P-Value	0.286	0.024	0.443	0.027

Robust standard errors are in parentheses and clustered by county. Baseline is an Ohio non-supermarket in February. Month coefficients and standard controls omitted.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.19: South Bend, IN Border

South Bend	(1) 2LCoke	(2) 12Coke	(3) 6Bud	(4) 24BudLite
Close	-0.53744***	-0.99392*	-1.24577***	-4.25793**
	(0.112)	(0.557)	(0.411)	(1.515)
Michigan	-0.23299	-3.88567***	-1.09247**	-1.31976
	(0.155)	(1.142)	(0.468)	(2.649)
Close*Michigan	0.56584***	2.80461**	1.94297***	2.53312
	(0.110)	(1.162)	(0.512)	(2.306)
Log Minutes*(1-Close)	-0.10490**	-0.36391*	-0.18162	-1.29484**
	(0.038)	(0.209)	(0.111)	(0.489)
Log Minutes*Michigan*(1-Close)	0.11212**	1.55739***	0.48381***	0.12064
	(0.043)	(0.394)	(0.166)	(0.830)
Log Med. HH Income	0.15217	0.36131	0.33815	-0.88157
	(0.206)	(0.608)	(0.367)	(1.602)
Log Density	0.00921	0.08945	0.08375	-0.02089
	(0.027)	(0.087)	(0.057)	(0.230)
Log Retail Density	-0.01457	-0.03329	0.01034	-0.05512
	(0.013)	(0.029)	(0.052)	(0.099)
Sale	-0.41084***	-0.90427***		-2.54815***
	(0.106)	(0.266)		(0.838)
Constant	0.51895	1.80313	2.64049	31.53890*
	(2.327)	(6.812)	(4.158)	(17.393)
Observations	53	45	38	45
R-squared	0.448	0.551	0.272	0.370
Border Gap Estimate	-0.333	1.081	-0.851	-1.213
P-Value	0.003	0.005	0.034	0.134

Robust standard errors are in parentheses and clustered by zip.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.20: Toledo, OH Border

Toledo	(1) 2LCoke	(2) 12Coke	(3) 6Bud	(4) 24BudLite
Close	0.75751*	-1.75340	-0.32586	-0.62753
	(0.364)	(1.793)	(0.304)	(0.688)
Michigan	2.88980***	-1.47089	0.55196	3.32186**
	(0.659)	(3.449)	(0.600)	(1.433)
Close*Michigan	-2.71886***	2.23654	-0.49293	-4.48952***
	(0.734)	(3.783)	(0.680)	(1.569)
Log Minutes*(1-Close)	0.33524**	-0.55614	-0.18373	-0.02293
	(0.145)	(0.753)	(0.137)	(0.309)
Log Minutes*Michigan*(1-Close)	-1.00568***	0.59504	-0.01945	-1.33795**
	(0.245)	(1.259)	(0.235)	(0.594)
Log Med. HH Income	0.06316	-0.37613	0.23825	0.79259**
	(0.205)	(0.601)	(0.179)	(0.297)
Log Density	0.02914	-0.16254	0.07568*	0.14583**
	(0.029)	(0.140)	(0.039)	(0.063)
Log Retail Density	0.00400	0.00500	-0.00849	0.02126***
	(0.004)	(0.010)	(0.006)	(0.006)
Sale	-0.41636***	-1.07084***		-0.65624*
	(0.063)	(0.182)		(0.328)
Constant	0.06808	11.37039	3.63489	8.64297**
	(2.421)	(8.453)	(2.145)	(3.870)
Observations	58	53	64	70
R-squared	0.446	0.336	0.295	0.282
Border Gap Estimate	-0.171	-0.766	-0.059	1.168
P-Value	0.198	0.073	0.697	0.001

Robust standard errors are in parentheses and clustered by zip.
*** p<0.01, ** p<0.05, * p<0.1

CHAPTER II

Equilibria and Location Choice in Corporate Tax Regimes: Implications of Preferential Taxation

2.1 Introduction

Continued globalization has resulted in a more integrated world market such that firms and capital investments have become increasingly mobile on an international scale. As a result, countries seeking additional tax revenues are implementing policies to attract these outside sources. One such policy is the preferential tax treatment of more mobile foreign firms and capital. Under a preferential regime, different tax bases have their own tax rates. Ireland is the most commonly cited example. The Finance Act of 1980 lowered the corporate income tax rate from 32% to 10% for the manufacturing and manufacturing related sectors. By the early 1990's, this lower rate had been extended to other sectors such as technology and finance. These preferential policies were targeted towards sectors with strong ties to foreign firms/investment, and are regarded as a primary reason for Ireland's high GDP growth.

However, the usage of such policies is frowned upon by policymakers in larger, multi-country organizations. Both the EU and OECD have suggested that the preferential treatment of these more mobile foreign bases is detrimental towards national tax revenues. Unilaterally, each country has a revenue incentive to adopt the less restrictive, preferential regime.[35] The adoption of preferential regimes by *all* countries, however, is believed to exacerbate tax competition in equilibrium and result in a race to the bottom. Competing preferential foreign rates could result in depressed national revenues due to the strategic interactions between countries. Conversely, banning preferential taxes could lessen tax competition as countries would only have access to a single, uniform tax rate for all bases. This debate therefore focuses on the difference in countries' equilibrium revenues when countries are allowed to adopt preferential policies versus when such policies are restricted. Are equilibrium tax revenues higher or lower when countries are competing with two, base-specific tax rates as opposed to one tax rate?[36]

Unlike policymakers, the economic literature is split on this issue. Janeba and Peters [1999] was one of the first papers to formally model this tax regime choice. The paper showed that a given country is weakly better off in revenue terms by allowing for multiple tax rates. In a Prisoner's Dilemma type argument, however, allowing all countries to adopt preferential policies results in lower global tax revenues. Keen [2001] found that the opposite is true. In a symmetric, two country model with two tax bases of differing mobilities/tax elasticities, a preferential regime focused tax competition on the relatively mobile sector. This shielded the relatively immobile sector and allows for greater overall revenue extraction. Conversely, a uniform tax rate exposes both sectors and results in lower revenues. In response, Janeba and Smart [2003] analyzed a more generalized model with aggregate tax base effects. The

[35] The set of possible tax rates under a preferential regime is larger, and includes the set of possible tax rates under the non-preferential regime. If a country decides to differentiate its tax rates and set a lower foreign, then it must be better off revenue-wise else it could have set a uniform rate.

[36] For the remainder of this paper, preferential or non-preferential will always refer to this larger, across all countries question.

Keen result could be characterized as a special case where the aggregate tax bases are fixed. In general, the authors found that a non-preferential regime is revenue-dominant.[37]

Other papers have added to and/or adjusted aspects of the basic models. Bucovetsky and Haufler [2007] found that Keen's result is robust to asymmetry of country sizes, while Haupt and Peters [2005] found that the opposite is true when adding home bias. Home bias is an empirically identified phenomenon whereby portfolios are skewed towards domestic investments - see Lewis [1999]. They found that partially restricting a preferential regime, i.e., limiting the rate gap between a country's tax rates, is revenue-improving. Oshima [2010] and Oshima [2009] separately introduced agglomeration and productivity asymmetry between countries.[38] He found that agglomeration supports the preferential regime while the effect of asymmetry is heavily dependent on the choice of production technologies. Wilson and Mongrain [forthcoming] allowed for heterogeneity in firm moving costs. The authors concluded that the non-preferential regime generates substantially more revenue when moving costs are distributed uniformly, while the opposite is true when there are relatively few firms with high moving costs.[39]

This paper advances the literature by analyzing a richer model that addresses three major issues with the current literature. First, most models assume that firms, investments, and/or countries are identical. This often leads to Bertrand style all-or-nothing responses. Equilibria are typically non-existent or, when they do exist, unrealistic. To construct models

[37] The literature on modeling tax competition falls into two camps. One camp follows the more classical conventions of Zodrow and Mieszkowski [1986] and Wilson [1986]. The other, New Economic Geography (NEG), follows the agglomeration models of Baldwin and Krugman [2004] in which firms generate positive spillovers by locating near each other. These NEG-style models typically result in some level of hysteresis as firms now have greater incentives to remain in larger clusters as opposed to moving for marginal tax charges.

[38] Burbidge et al. [2006] also considers differential cross-country productivity. Their emphasis, however, is on the shares of foreign ownership in each firm/country.

[39] Surveys of the theoretical tax competition literature include Genschel and Schwarz [2011], and Wilson [1999]. Devereux et al. [2008] provide a review of recent empirical trends in corporate taxation.

with more realistic interior solutions, some form of heterogeneity must be present in the model. Wilson and Mongrain [forthcoming] and Baldwin and Krugman [2004] do so by having heterogeneous moving costs and agglomeration technologies, respectively. Another common method is simply to assume smoothness over the base/revenue functions. Second, firms' potential domestic and foreign profits/returns are assumed to be equal, i.e., gross profits are independent of location. I refer to this as the property of perfectly correlated (one-to-one) cross-country profits. This also has the property of inducing the all-or-nothing response as it assumes that countries are identical with regards to profit generation. Third, while previous papers have expanded models to include aspects such as moving costs and asymmetric countries, they have all done so separately. The literature currently lacks a comprehensive model that jointly incorporates these considerations. Papers have relied on relatively simplistic models due to the analytical intractability of solving for equilibrium tax rates and revenues in less restrictive settings.

In this paper, regime choice is analyzed in the context of a single period, two country model which jointly incorporates the following factors. First, firms are heterogeneous in their profits. Second, countries are allowed to be asymmetric in both population (the number of firms) and productivity. Third, moving costs and moving cost deductibility are introduced. Fourth, I consider the two cases of perfectly correlated and non-correlated profits. The goal of this paper is to analyze the revenue implications of preferential versus non-preferential regimes. Specifically, are there cases where countries can generate greater revenues in an equilibrium with preferential regimes? Given the introduction of imperfect profit correlation, we now have the possibility for "natural" firm movement in the absence of taxes.[40] Therefore, I also examine the impact of regime choice on firms' location choices. This is quantified by

[40]This is novel in the tax competition literature and represents a significant divergence. Any cost-inducing movement was previously assumed to be inefficient.

comparing the level of total, net-of-moving-cost profits with and without taxes.[41]

As expected, the analytical results from this model are limited, so a numerical, computational approach is used to construct countries' best response functions and the resulting pure strategy Nash Equilibrium revenues. Four main results are presented in this paper. First, the preferential regime generates more revenue than the non-preferential regime in cases where productivities are sufficiently asymmetric and profits are non-correlated. These country-level differences mitigate the negative effects of tax competition by decreasing the relative number of tax-sensitive firms. Because countries are effectively targeting a distinct set of firms within each base, preferential taxation generates greater equilibrium revenues. This result is strengthened if the more productive country has less population. Higher moving costs increase (decrease) this likelihood when a country is more (less) productive, while higher moving cost deductibility tends to go against the preferential regime. Second, the choice of regime does not matter when moving costs are zero. Under perfect mobility, both regimes generate identical equilibria. This result is robust to profit correlation and country asymmetries. Third, pure strategy Nash Equilibria do not exist when profits are perfectly correlated and moving costs are strictly positive. This is caused by discontinuities in countries' best response functions. Fourth, the preferential regime impacts firms' location choices to a lesser extent than the non-preferential regime except when moving cost deductibility is high.

The remainder of the paper proceeds as follows. Section 2 lays out the analytical model. Section 3 details the numerical model. Section 4 presents the four main results in greater detail. Finally, Section 5 concludes and discusses policy implications.

[41]Section 3 discusses how we can connect location choice to traditional connotations of efficiency.

2.2 The Standard Model

Following the notation of Keen [2001], consider a single period game with two countries and two tax bases. The size of each base is normalized to one so let b and B denote the fractions of each base that locate in Country A. Country A has access to tax rates t and T which are the rates applicable to bases b and B, respectively. The other country's tax rates are denoted by t^* and T^*.

Because governments are assumed to be revenue maximizers, the objective function for Country A is given by:[42]

$$\max_{t,T} R_A = tb(t,t^*) + TB(T,T^*) \tag{2.2.1}$$

Under a preferential regime, t and T are allowed to be different so countries maximize with respect to each rate. Expressions for Country A's optimal tax rates, assuming convexity of the base functions, are given by:[43]

$$b + tb_t = 0 \quad \rightarrow \quad \overbrace{b_t \frac{t}{b}}^{\varepsilon_t^b} = -1 \tag{2.2.2}$$

$$B + TB_T = 0 \quad \rightarrow \quad \underbrace{B_T \frac{T}{B}}_{\varepsilon_T^B} = -1 \tag{2.2.3}$$

From the left-hand side of (2.2.2), we can see that Country A's optimal t is found by equating the marginal benefit and marginal cost. Intuitively, a small increase in A's rate

[42]Maximizing revenue is arguably the true objective function for policymakers as indicated by EU/OECD reports. From a modeling standpoint, maximizing revenue also requires fewer assumptions than welfare maximization. Additionally, the results will be equivalent if we assume that the social marginal benefit of a dollar of tax revenue is higher than a dollar held privately.

[43]These require that the base functions be smooth, differentiable, and quasi-concave.

increases revenues collected from its current base, b, at the cost of lost revenue from those that exit, tb_t. Likewise, we have the analogous interpretation of the optimal T in (2.2.3). Alternatively, we can express the optimal conditions in terms of tax elasticities. This gives us the standard, consumer pricing result which says that revenue is maximized by selecting the price (tax rate) at which demand (tax base) is unit elastic. If Country A adopts the preferential regime, tax rates are chosen separately such that both elasticities are unit elastic. If it adopts the non-preferential regime, then the two rates must be equal and jointly considered as seen in (2.2.4). We can therefore denote both bases as being functions of the uniform rates τ and τ^*. Notice that the uniform tax rate affects both bases simultaneously. A marginal increase in the uniform rate now generates greater, concurrent gains and losses from both bases.

$$b + \tau b_\tau + B + \tau B_\tau = 0 \rightarrow b(1 + \underbrace{\frac{\tau}{b} b_\tau}_{\varepsilon_\tau^b}) = -B(1 + \underbrace{\frac{\tau}{B} B_\tau}_{\varepsilon_\tau^B}) \tag{2.2.4}$$

What can we infer from these equations? If the bases are identical and/or have equal tax elasticities, then there are no differences between the two regimes, i.e., tax rates and revenues are the same regardless of whether Country A optimizes over the two bases jointly or separately.[44] From a third degree price discrimination standpoint, this is equivalent to the optimal pricing problem when there are two equal demand groups. Group specific pricing is only revenue enhancing if the groups are different. Otherwise, the group specific rates are equal and a single, uniform price is sufficient. More generally, the rates will be same if the elasticities are unit elastic at the same, optimal tax rate.

We can also generalize the model to include population. In the tax competition literature,

[44]This basic result has been noted in Keen [2001] as well as a number of other papers.

population refers to the sizes of the bases. Keeping the masses of the bases normalized to one, let N_b and N_B be population multipliers such that Country A now receives revenues of $tN_b b + TN_B B$. Notice that the population multipliers do not affect the optimal preferential tax rates as they drop out of (2.2.2) and (2.2.3). Population does, however, affect the uniform rate. We can think of the population ratio, N_b/N_B, as the relative weight placed on revenues from b as opposed to B. The optimum uniform tax rate will therefore be the same for equal N_b/N_B ratios such as $(N_b, N_B) = (2, 1)$ and $(N_b, N_B) = (4, 2)$. Taking the tax rates of the other country as a given, we can also say that $t \geq \tau \geq T$ will be true if base b is weakly less mobile/elastic than base B.[45]

As shown in Janeba and Peters [1999], each country has a weak incentive to unilaterally adopt the preferential regime. Recall that the non-preferential regime is a special case of the preferential regime where the two rates are constrained to be equal. Therefore, preferential taxation is a weakly dominant strategy, i.e., revenues are equal if not greater with two separate tax rates regardless of the other country's choice of rates and/or regime. However, the resulting equilibrium revenues when *both* countries adopt preferential regimes may or may not be higher than revenues when *both* countries have non-preferential regimes.

Evaluating revenues when the base elasticities are different is problematic. First, we cannot derive a specific solution for the optimal tax rate as is oftentimes true in optimal taxation problems. Even if we applied a functional form to b and B, it would not always be possible to isolate for the tax rates let alone characterize the *equilibrium* tax rates. Second, recall that the central research question is to compare equilibrium revenues across regimes. This is difficult due to the difference in the number of choice variables, i.e., tax instruments. Previous papers have worked around this issue by making simplifying assumptions to reduce

[45] The proof is shown in the Appendix C. The opposite result will hold if B is less responsive than b.

the number of choice variables in the preferential regime's optimization problem.

Keen [2001] does so by making three assumptions. First, he assumes that the bases depend only on the cross-country rate differentials, $t - t^*$ and $T - T^*$, such that the actual tax rate levels no longer matter. Second, the bases are assumed to be smooth with respect to the differentials. Third, the two countries and the resulting equilibrium are assumed to be symmetric. Together, these three assumptions allow Keen to say that the rate differentials must be equal to zero in equilibrium. Because the bases only depend on the differential, the actual tax rates become irrelevant and the equilibrium revenues are now only functions of $b(t, t^*)$ and $B(T, T^*)$ for any set of tax rates where $t - t^* = 0$ and $T - T^* = 0$. This allows for comparisons across regimes. Janeba and Smart [2003] reduce the variable space by exogenously fixing the within-country rate gap, $t - T$ and $t^* - T^*$, to be some constant θ. This is meant to represent a small restriction on the preferential regime. It is an approximation of the unrestricted preferential regime if we assume that the actual rate gap is close or equal to θ. The strategies employed by the two previous models highlight the need for additional assumptions in order to gain analytical tractability. Below, I describe the numerical model which moves away from such assumptions.

2.3 The Numerical Model

Consider again the single period model with two countries, A and B. In each country, there is a continuum of profit maximizing, atomistic firms. Alternatively, we can include multinationals by interpreting these larger firms as a collection of atoms where each individual atom would then represent a separate activity, operation, or subsidiary. We can also interpret these atoms as units of capital investment which can be invested domestically or abroad in the form of FDI. I will use the firm interpretation for the remainder of the paper.

Each firm is characterized by its profit pair, (π_A, π_B), where π_j denotes the profits it would make if located in country j.[46] This is private information known by the firms. Conversely, governments only know that the distribution of firms/profit pairs for those originally located in country j follows $f_j(\pi_A, \pi_B)$. For the numerical model, assume that both the domestic and potential foreign profits of original firms in country j are distributed uniformly in 0 and $\bar{\pi}_j$. Under the assumption that $f_A = f_B$, this implies that the profit pairs in both countries are drawn from a bivariate uniform distribution over $[0, \bar{\pi}_A]$ x $[0, \bar{\pi}_B]$. Additionally, we can say that Country A is more productive than Country B if $\bar{\pi}_A > \bar{\pi}_B$.[47] This allows us to analyze symmetric productivity cases such as $(\bar{\pi}_A, \bar{\pi}_B) = (1, 1)$, as well as asymmetric cases when the upper bounds are different. As in the previous section, country size/population is defined as the number of firms originally in a country and is denoted by the multipliers N_A and N_B. The different combinations of productivity and population, e.g., $(\bar{\pi}_A, \bar{\pi}_B, N_A, N_B) = (1, 1, 2, 1)$, allow us to simulate a number of different scenarios.

The timing of the model proceeds as follows. Countries are initially endowed with a stock of original firms. Governments simultaneously set profit taxes t_j and T_j knowing only the distributional characteristics and population of firms. t_j applies to j's domestic firms, those originally from j that remain in j, while T_j applies to its foreign firms, those originally from k that move to j. Tax rates are bounded between 0 and 1 (0% and 100%) in increments of 0.001 (0.1%). After taxes are set, firms receive their profit draws and make their location choices under full information. Firms can choose to locate domestically or in the foreign

[46] These profits can be generated under a monopolistic competition setting similar to a simplified Melitz-style model. Each atom has a unit of capital that it can locate domestically or abroad. Production is given by $q_{ij}(k) = \varphi_{ij}k$ where φ_{ij} is a firm-country specific productivity draw. Assume that greater productivity always generates greater firm profits, i.e., demand is always inelastic. Alternatively, we can assume that prices are constant (above marginal cost) across countries in which case productivity directly translates into profits.

[47] While this restriction on identical f's seems strong, empirical work has suggested that firms, at least in OECD countries, are drawn from identical Pareto distributions.

country by incurring a fixed moving cost c. The c term can incorporate the monetary cost of physically relocating and/or adjusting to new regulations or practices in the foreign country. Finally, profits and tax revenues are realized.

One concern is that models with discrete tax steps may generate different equilibria than those with continuous tax rates. For example, consider the standard Bertrand model with two identical zero cost firms. Allowing for ε undercutting results in a single pure strategy Nash Equilibrium where prices are zero for both firms. If prices are restricted to be in one cent increments, however, this generates an additional pure strategy equilibrium where both firms set a price of $0.01 and receive positive profits. However, discrete tax steps are arguably more appropriate in this model. Actual tax rates are fairly discrete and rarely move into the hundredth or thousandth decimal. Looking at a model where rates are continuous may be unrealistic and incorrect if equilibria arise solely due to this construction.

Because firms know their profit draws and tax rates ex ante, location choice is decided based on the highest net-of-tax, net-of-moving-cost profits. Therefore, a firm originally in country j will move to country k if:

$$\underbrace{\pi_k}_{\text{profit in } k} - \underbrace{T_k(\pi_k - \alpha c)}_{\text{tax payment net deductions}} - c > \underbrace{\pi_j}_{\text{profit in } j} - \underbrace{t_j \pi_j}_{\text{tax payment}} \qquad (2.3.1)$$

where α is the fraction of the moving cost deductible from a firm's taxable liabilities.[48] Following the notion of home bias, assume that firms always choose to stay in the domestic country when (2.3.1) holds with equality.

The government's objective is to maximize tax revenues by choosing the optimal t_j and

[48] Negative tax liability is allowed in the model for cases where moving cost deductions exceed profits. This simulates any potential benefit from tax credit carryover. However, this never occurs as moving would always be sub-optimal in these situations. As such, the results of the model are unaffected by this assumption.

T_j. Because this paper focuses on the choice of rates, assume that α is taken as given. Total tax revenue in country j, R_j, is given by the sum of revenue from domestic firms, $R_j^D(t_j, T_k, \alpha, f)$, and foreign firms, $R_j^F(t_k, T_j, \alpha, f)$.

$$R_j^D = N_j \int_0^{\bar{\pi}_j} \int_0^{\frac{(1-t_j)\pi_j + (1-\alpha T_k)c}{1-T_k}} t_j \pi_j f \, \mathrm{d}\pi_k \, \mathrm{d}\pi_j \tag{2.3.2}$$

$$R_j^F = N_k \int_0^{\bar{\pi}_j} \int_0^{\frac{(1-T_j)\pi_j - (1-\alpha T_j)c}{1-t_k}} T_j(\pi_j - \alpha c) f \, \mathrm{d}\pi_k \, \mathrm{d}\pi_j \tag{2.3.3}$$

Assuming that the revenue functions are continuous, differentiable, and quasi-concave, we can maximize with respect to the tax rates. Under most distributions, however, we cannot solve for the exact tax rates.[49] Under the bivariate uniform distribution, the revenue function is also highly piece-wise. I therefore analyze the model using MATLAB to manually construct countries' revenue and best response functions. Pure strategy Nash Equilibria are then calculated accordingly. Further detail on the optimum rates, coding procedure, and range of parameters/cases examined are given in Appendix F.

In terms of profit correlation, I consider perfectly correlated and non-correlated profits. Under non-correlated profits, a firm's potential profit from locating abroad is independent of its domestic profit. This implies that $f(x, x) = f(x, y)$ for all $x \neq y$ in the range of π. The distribution of firms originally in country j can therefore be represented by a uniform rectangle with π_A and π_B on the axes. Under perfectly correlated profits, firms make the

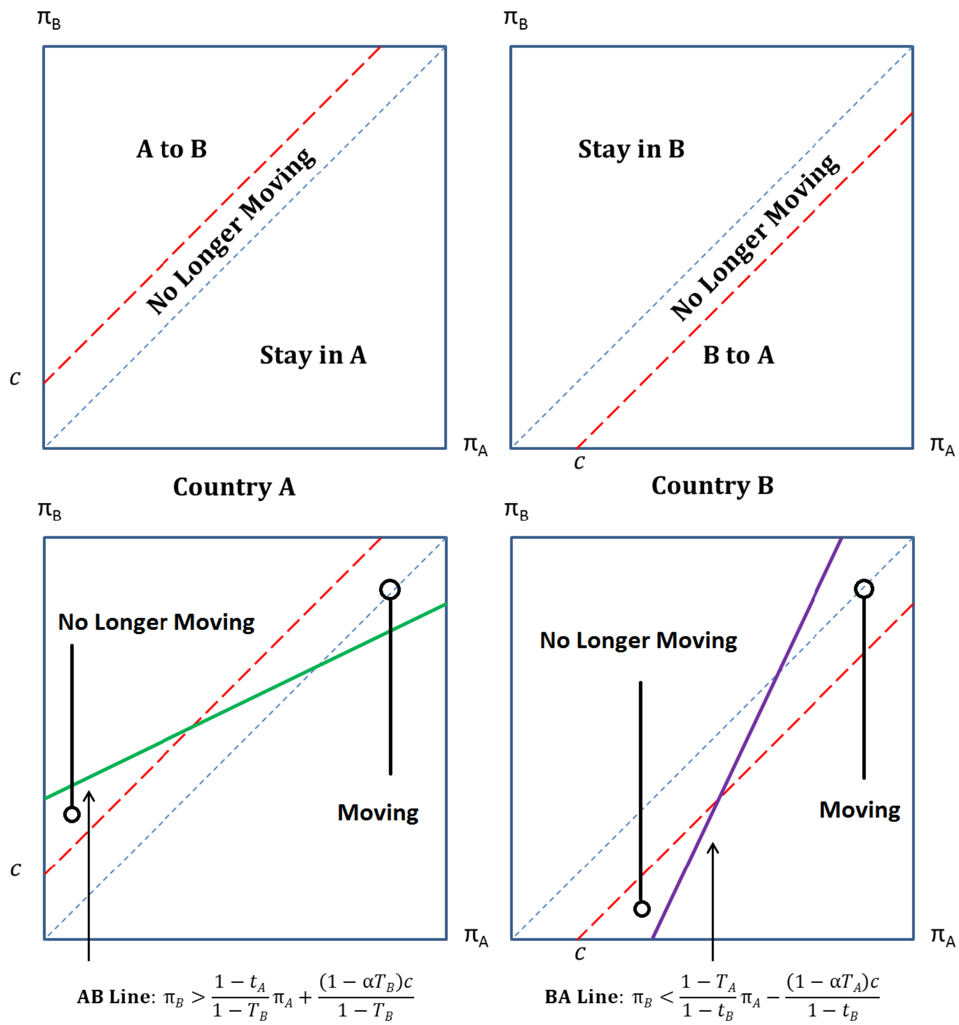
[49] While closed form solutions may exist, the tax variables are non-separable. See Appendix sections D and E for the optimal tax rate formulas under the previous assumptions.

same profit in both countries. This implies that $f(x, y) = 0$ for $x \neq y$ such that positive density exists only on the 45° line with support $[0, \bar{\pi}_j]$. Fundamentally, profit correlation can be interpreted as a property of the f distribution. For a given f, profit correlation is greater the faster the density function falls when moving away from the diagonal because profit pairs closer to $\pi_A = \pi_B$ are receiving relatively greater weight. Under perfectly correlated profits, the density falls immediately to zero for any $\pi_A \neq \pi_B$. Under non-correlated profits, uniformity implies that all profit pairs have equal probabilities such the density function is flat and does not fall.[50]

Figure 2.3.1 depicts the distribution of firms in the case of non-correlated profits with $(\bar{\pi}_A, \bar{\pi}_B) = (1, 1)$. Notice that population does not affect the graphical interpretation. Firms located on the dotted, 45° line are equally profitable domestically and abroad. Those to the

[50]The level of profit correlation can also proxy for the costliness of profit shifting between high and low tax jurisdictions. If firms can freely shift profits, this is equivalent to assuming that profits are location-independent, i.e., profits are perfectly correlated. As profit shifting becomes more costly, this is equivalent to a decrease in profit correlation.

Figure 2.3.1: Distribution of Firms - $(\bar{\pi}_A, \bar{\pi}_B) = (1, 1)$



left are more profitable in Country B while those to the right are more profitable in A. The two dashed lines indicate the addition of moving costs. Now, only those firms that make more profit net of c will move. This results in the two indicated wedges in the top two graphs. The two solid lines in the bottom two graphs represent the net-of-taxes, net-of-moving cost equations given by (2.3.1) for an arbitrary set of tax rates. I will refer to these as the AB and BA lines, i.e., the A to B and B to A moving lines. Those firms originally in A that are located left of the AB line will move to B, while those to the right will remain in A, and analogously for firms in Country B and the BA line. In comparison the pre-tax movement of firms in the top graphs, the introduction of taxes and moving cost deductibility now creates the four additional wedges in the bottom graphs.

When taxes are absent, firms move to the location that affords them the highest net-of-moving cost profit. The presence of tax rates potentially alters this choice. The degree to which each regime distorts location choice is quantified by the total net-of-moving cost profit (TNP) loss. TNP for a given set of tax rates is equal to the sum of all firms' profits less c. The TNP loss is therefore the difference in TNP between the no-tax case and the with-tax case. Under the firm movement interpretation, we can connect location choice distortions (TNP loss) with traditional notions of efficiency (deadweight loss) if we assume that higher profits are driven by greater productivity/lower costs, e.g., locating near a required natural resource. However, moving for higher profits can be inefficient if the profit gains are driven by greater market power.[51] Graphically, we can see the tax distortions by comparing the AB and BA lines (movement with taxes) against the dashed moving cost lines (without taxes) in Figure 2.3.1. When all tax rates are zero, the two sets of lines coincide and there are no distortions.

A number of basic results can be seen from this figure. When moving costs are zero, all equilibria where tax rates are harmonized, $t_j = T_k$ and $T_j = t_k$, generate zero TNP loss. When moving costs are positive, any positive tax rates will cause the two sets of lines to diverge so long as deductibility is not 100%. Under non-correlated profits, this guarantees some amount of TNP loss. Under perfectly correlated profits, positive equilibrium rates can

[51] Traditionally, the issue of tax competition and efficiency has centered around two main concepts. First, if tax revenues are applied towards public good provision, then tax competition typically results in inefficiently lower rates/revenues. Alternatively, if government is seen as a leviathan which sets tax rates that are too high, then tax competition is a means to check this power and may be efficiency-improving.

still be non-distortionary as long as the AB and BA lines do not intersect the diagonals line, i.e., when the domestic rates are weakly smaller than the competing foreign rates. Recall that there are no firms located in the off-diagonal areas. When $c = 0$, all allocations result in zero TNP loss under perfectly correlated profits. This is not true under non-correlated profits due to the fact that some firms will naturally move. As moving costs increase, however, both the number of movers and the no-tax TNP will decrease.

2.4 Results

The results center around the revenue and loss gaps. The revenue gap for country j is defined as the difference between j's equilibrium revenue when countries are using preferential regimes and the equilibrium revenue when countries are restricted to using non-preferential regimes. Therefore, a positive revenue gap implies that country j does better revenue-wise under the global preferential regime. A negative revenue gap implies that the non-preferential regime is revenue-dominant. Likewise, the loss gap is defined as the differential TNP loss between the two regimes. A positive loss gap implies that the preferential regime is *less* distortionary because it generates a lower TNP loss.

I will present the main results of the paper by starting with the simplest/most restrictive model and comparing the changes when elements are added/relaxed.

2.4.1 Perfect Mobility and Regime Indifference

In the simplest and most restrictive baseline model, assume that profits are perfectly correlated across countries. Under perfect mobility, both the preferential and non-preferential regimes result in the same equilibrium rates and revenues. This implies that the equilibrium domestic and foreign rates are the same, and also equal to the equilibrium uniform rate. The revenue gap is therefore zero for both countries. Likewise for location choice, both regimes induce the same level of TNP loss. This result holds even if we relax assumptions on profit correlation and/or symmetric countries.

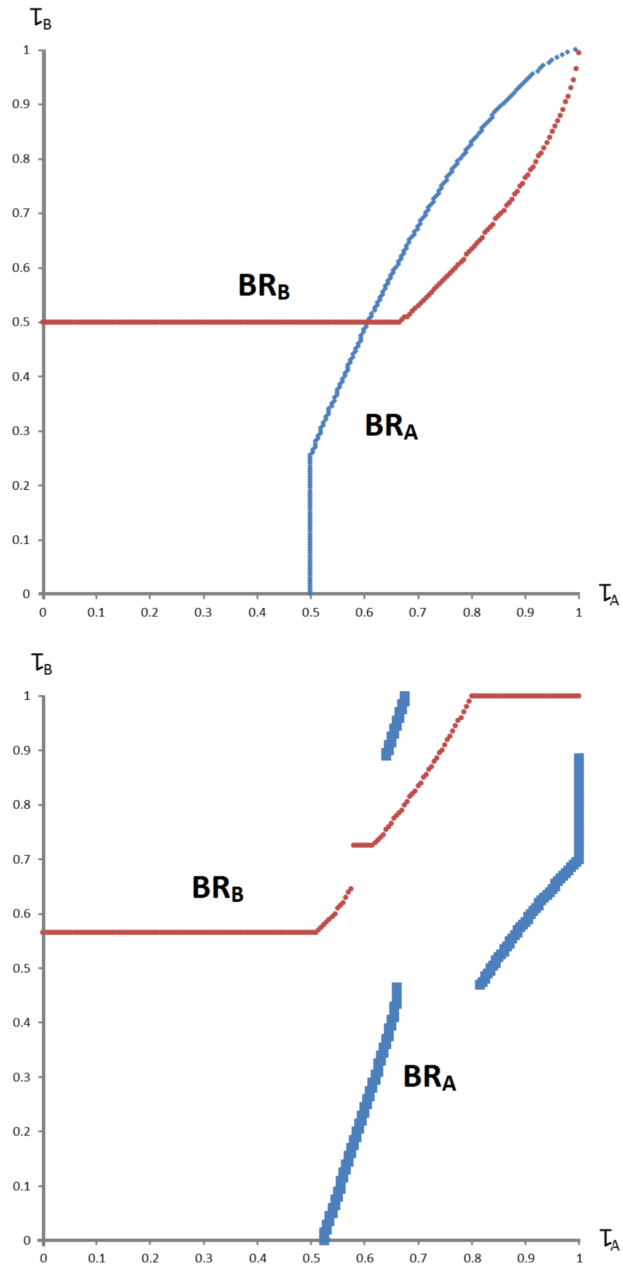
Proposition 1. *Under perfect mobility, the revenue and loss gaps are both zero. The choice of regime does not matter when $c = 0$.*

Under perfect mobility, the distinction between the bases disappears. Being originally from Country A or Country B no longer matters if firms can move costlessly. As previously discussed, this result can be interpreted in the form of a third degree price discrimination problem with two equally responsive demand groups. The optimum group-specific prices are the same which implies that the optimum uniform price will also be the same. This is true when moving costs are zero and distributions are the same across countries.

As an example, consider the case of $(\bar{\pi}_A, \bar{\pi}_B, N_A, N_B) = (1.5, 1, 1, 1)$ under perfect mobility and non-correlated profits. Deductibility does not matter for $c = 0$. While both countries have the same population, A is more productive than B. This implies that for every firm in B, there exists a corresponding firm in A that generates 50% more profit. This also implies that two-thirds of the firms originally in B would generate more gross profit if they moved to A. The top graph of Figure 2.4.1 depicts the best response functions for A and B under the non-preferential regime. The intersection indicates a unique pure strategy equilibrium at $(\tau_A, \tau_B) = (0.609, 0.5)$ which results in revenues of $(R_A, R_B) = (0.69219, 0.28417)$.[52] As

[52] The average statutory corporate tax rate in the real world is around 23% which is lower than most rates in this model. This higher simulated rate is a result of the uniform distribution. Niu [2013a] simulates a numerical model under a bivariate Pareto distribution of profits which generates more realistic tax rates.

Figure 2.4.1: Best Responses - $(1.5, 1, 1, 1)$ $c = \{0, 0.3\}$



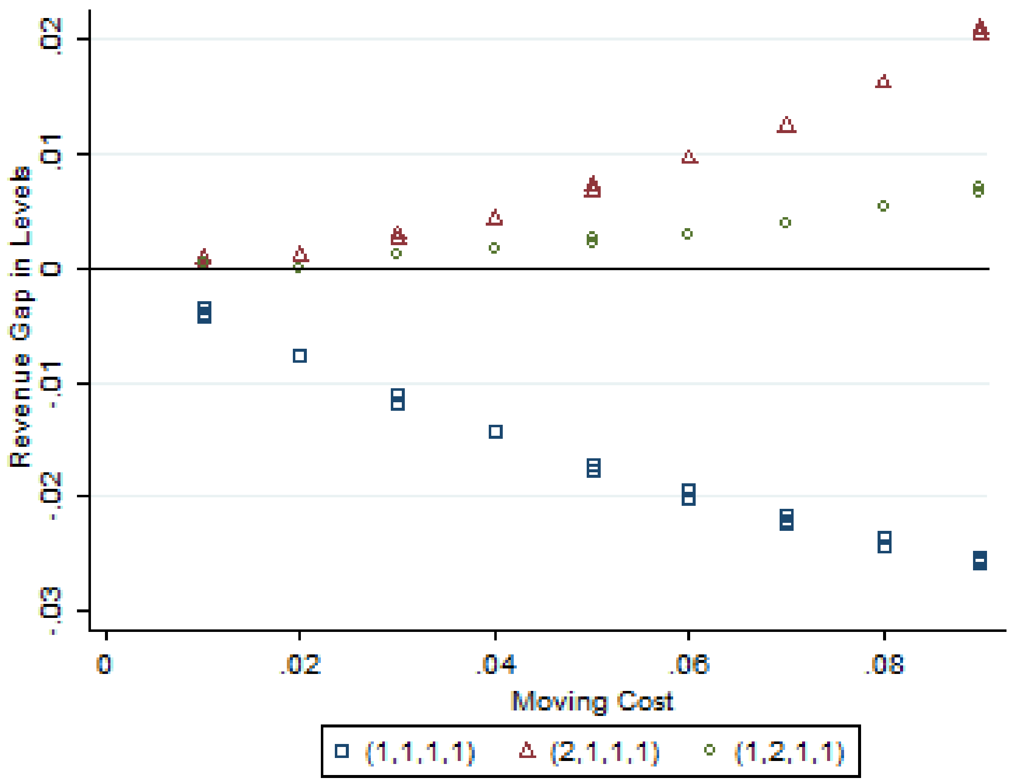
expected, A's rate is higher because a greater proportion of firms are more profitable in A than B. This makes it easier for A to both keep and attract firms, and provides it greater revenues in equilibrium. Regarding location choice, the non-preferential regime generates a TNP loss of 0.01727 which is roughly 1% of the no-tax TNP. Under the preferential regime, A's equilibrium domestic and foreign tax rates are both 0.609. Likewise, B's equilibrium rates are both 0.5. This results in identical equilibrium revenues and TNP losses. Therefore, the revenue and loss gaps are both zero.

In cases where countries are equally productive, the two regimes can generate unequal numbers of equilibria. For the case of $(\bar{\pi}_A, \bar{\pi}_B, N_A, N_B) = (1, 1, 1, 1)$, the non-preferential regime generates two symmetric equilibria with uniform rates of $\tau = \{0.5, 0.501\}$ and per country revenues of $R = \{0.3333, 0.334\}$. The preferential regime, however, generates four equilibria. Two of them correspond to the non-preferential equilibria. The other two equilibria have combinations of domestic and foreign rates at 0.5 and 0.501 which result in per country revenues of 0.33367. Because we cannot justify the likelihood of one equilibrium over the other, a comparison can only be made on the range of equilibrium revenues.[53] TNP losses are equal to zero in all four preferential equilibria.

As a related result, Figure 2.4.2 plots the revenue gap for a number of different cases in the neighborhood of $c = 0$. What this suggests is that the revenue gap approaches zero as moving costs approach zero. While this monotonic relationship does not always hold at higher values of c, it does imply that decreasing moving costs eventually reduce the potential revenue losses from having the "wrong" regime.

[53] If we compare each of the corresponding equilibria in cases where $c = 0$, the mean revenue gap is roughly 0.00004 or 0.009% in percentage terms.

Figure 2.4.2: Moving Costs and the Revenue Gap Near $c = 0$



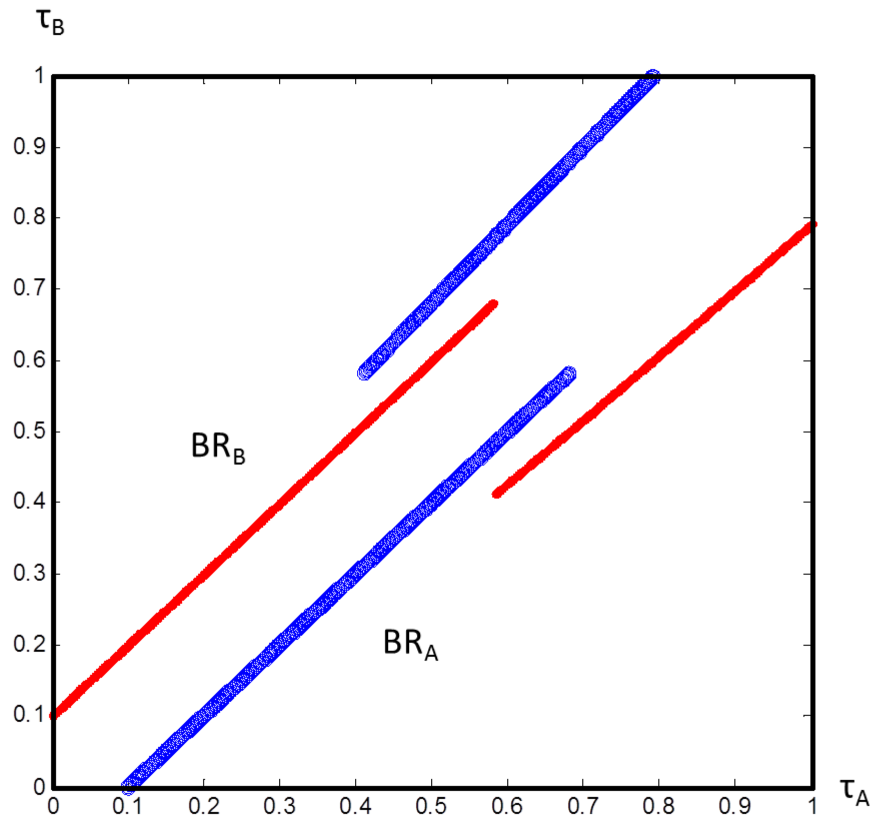
2.4.2 Perfect Correlation and Non-Existence

Because zero moving costs result in zero revenue and loss gaps, assume that moving costs are now strictly positive. Returning to the baseline model with perfectly correlated profits, symmetric countries, and non-zero values of c, I find that the two regimes are again identical when moving costs are very high. When c is above a max taxation threshold, firms become sufficiently immobile that countries are now able to fully tax away profits. Further analysis will therefore consider levels of c greater than zero but below this max taxation threshold. Under the non-preferential regime, however, pure strategy equilibria no longer exist in this range of moving costs. This is caused by the perfect profit correlation assumption.

Proposition 2. *When profits are perfectly correlated and moving costs are positive, the non-preferential regime does not generate any pure strategy equilibria.*

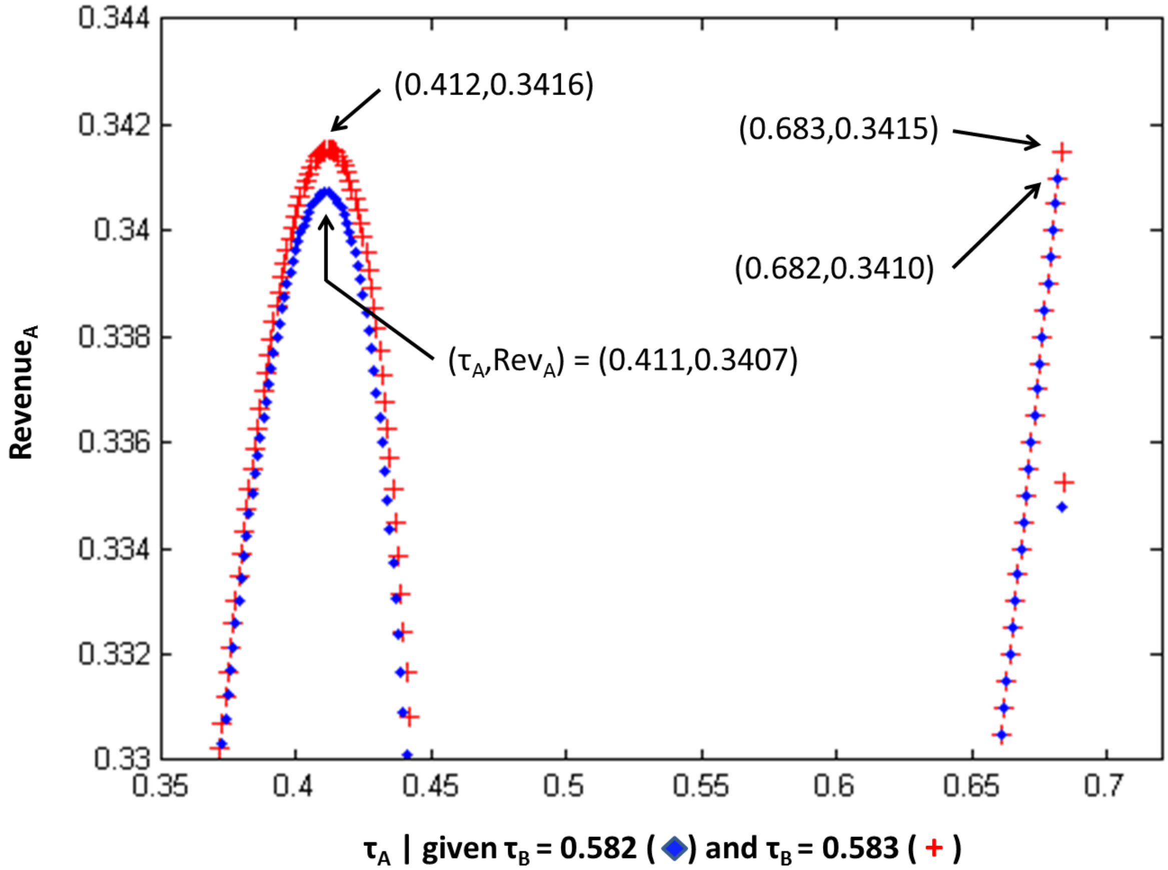
We can see this non-existence in Figure 2.4.3 which depicts the non-preferential best response functions for the case of $(1, 1, 1, 1)$ and $c = 0.1$. Notice that an intersection does not exist due to the discontinuities in both functions. From an iterative standpoint, Country A responds to Country B's high uniform rate by undercutting. A small, incremental decrease in the uniform tax rate results in large changes to the domestic and foreign bases. Country B's best response is to then undercut A's uniform rate. Under perfect mobility, this continues until tax rates reach 0.001 or 0. Notice that setting uniform rates of zero always constitutes an equilibrium. For any given tax rate, the best response is to always undercut when firms are perfectly mobile. However, this is no longer true with positive moving costs as countries now have an immobile group of domestic firms. In the Bertrand literature, this is similar to the concept of a loyal base. Given these immobile firms, a uniform rate of zero is never optimal as both countries can simply tax their loyal bases to generate positive revenues. Instead of continuously undercutting at very low tax rates, countries can do better by setting a higher tax rate to exploit these immobile firms. We therefore have this jump/discontinuity phenomenon in the best responses.

Figure 2.4.3: Best Responses - $(1, 1, 1, 1)$ $c = 0.1$



Returning to Figure 2.4.3, we can see that A finds it optimal to undercut when B's tax rates are high. Undercutting pushes both countries' rates lower and lower until they approach the discontinuity. When $\tau_B = 0.583$, A's best response is to undercut and set a tax rate of $\tau_A = 0.412$. When B's tax rate drops to $\tau_B = 0.582$, however, A now finds it optimal to stop undercutting and instead, jumps up to set a higher tax rate of $\tau_A = 0.682$. Figure 2.4.4 plots the revenue function for A around the discontinuity. With A setting a uniform rate of $\tau_A = 0.682$, B finds it optimal to undercut and the cycle restarts. Thus, there is no stable pure strategy equilibrium.

Figure 2.4.4: Revenue Graphs for Country A - Before and After the Jump



Notice that this only occurs under the non-preferential regime. Recall that a single, uniform rate must simultaneously balance revenue considerations over both the domestic and foreign bases. By jumping up, the uniform rate switches from focusing on both bases to focusing more on its immobile, domestic base. Intuitively, it concedes some level of competition for foreign firms so it can extract higher revenues from its domestic firms. Because non-existence, i.e., the jump phenomenon, occurs only because of this switch in focus, existence is not an issue under the preferential regime. With separate taxes for each base, countries do not have to switch focus and can undercut each base separately. For example, the preferential regime generates equilibrium rates of $(t_A, T_A, t_B, T_B) = (0.1, 0, 0.1, 0)$ with revenues of $(R_A, R_B) = (0.5, 0.5)$ in the previous example.

Non-existence is a common issue in this literature.[54] While some papers introduce functional assumptions to guarantee existence, other papers such as Davies and Eckel [2010] appeal to mixed-strategies when pure strategies fail. However, the real world applicability of mixed strategies is questionable. In actuality, tax rates do not appear to vary from year to year. Alternatively, appealing to a sequential, Stackelberg style equilibrium has been proposed by Janeba and Peters [1999] and Mongrain et al. [2010]. It is also possible to show the existence of equilibrium loops where countries cycle through a closed set of tax rates. This paper remains agnostic on the issue so subsequent results focus solely on cases where pure strategy equilibria exist.

Non-existence also occurs when profits are non-correlated but to a much lesser extent. This only occurs for higher values of c because the jump phenomenon requires a relatively large proportion of very immobile firms. As an example, the lower graph in Figure 2.4.1 from the previous result depicts the best response functions for $c = 0.3$.

2.4.3 Preferential Regime and Sufficient Asymmetry

Returning to the baseline model, we now consider a model with symmetric countries, non-correlated profits, and non-zero moving costs. In these cases, the revenue gap is negative, i.e., the non-preferential regime is revenue-dominant. In the case of $(1, 1, 1, 1)$ and $c = 0.1$, the revenue gap is -0.026 which represents a difference equivalent to 7.1% of the non-preferential

[54] This is also an issue of contention in the game theory literature. Börgers [1992] suggests using iterated deletion of strictly dominated strategies in lieu of mixed strategies.

revenue. Allowing for asymmetric productivities and populations, however, flips the revenue gap.

Proposition 3. *The preferential regime is revenue-dominant if productivities are sufficiently asymmetric. This result is strengthened if the more productive country is also less populous.*

Figure 2.4.5 plots the revenue gap for countries with equal populations but productivity asymmetries that vary between 1% and 50%. Specifically, I fix the productivity of Country B at $\bar{\pi}_B = 1$ but consider productivities for Country A in the range of $\bar{\pi}_A \in [1.01, 1.5]$.

Figure 2.4.5: Revenue Gap and Productivity Asymmetry I

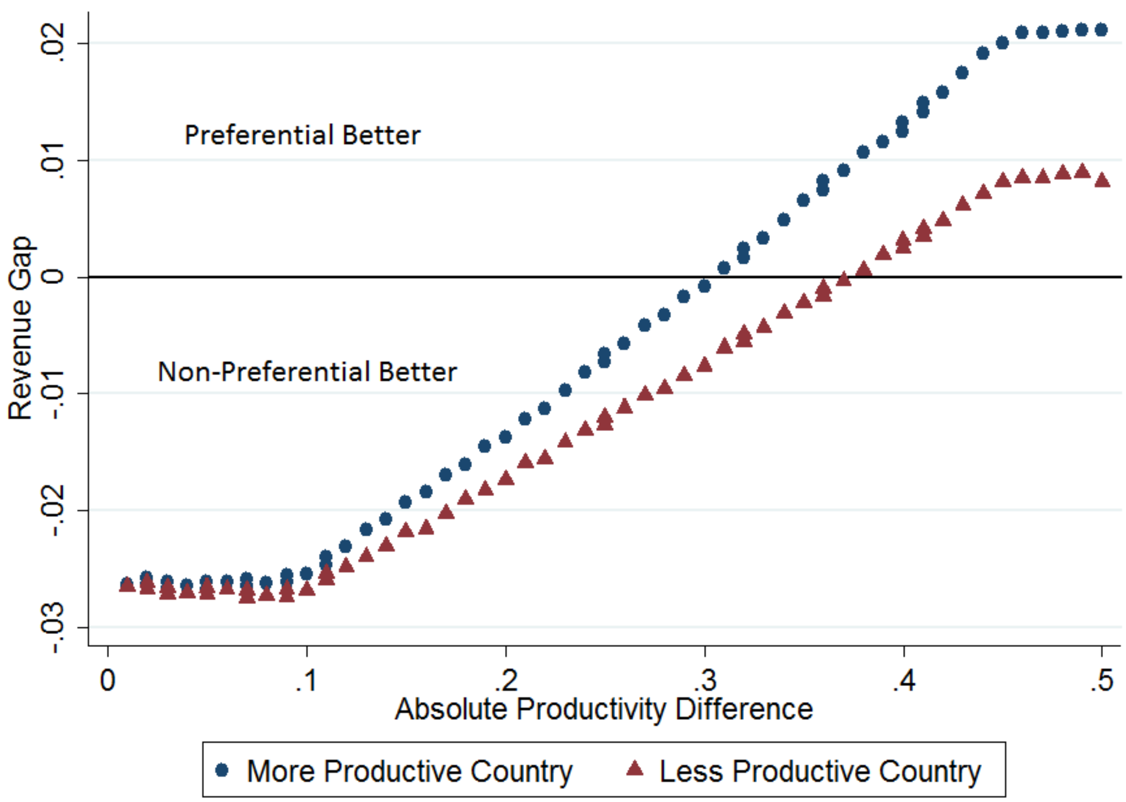
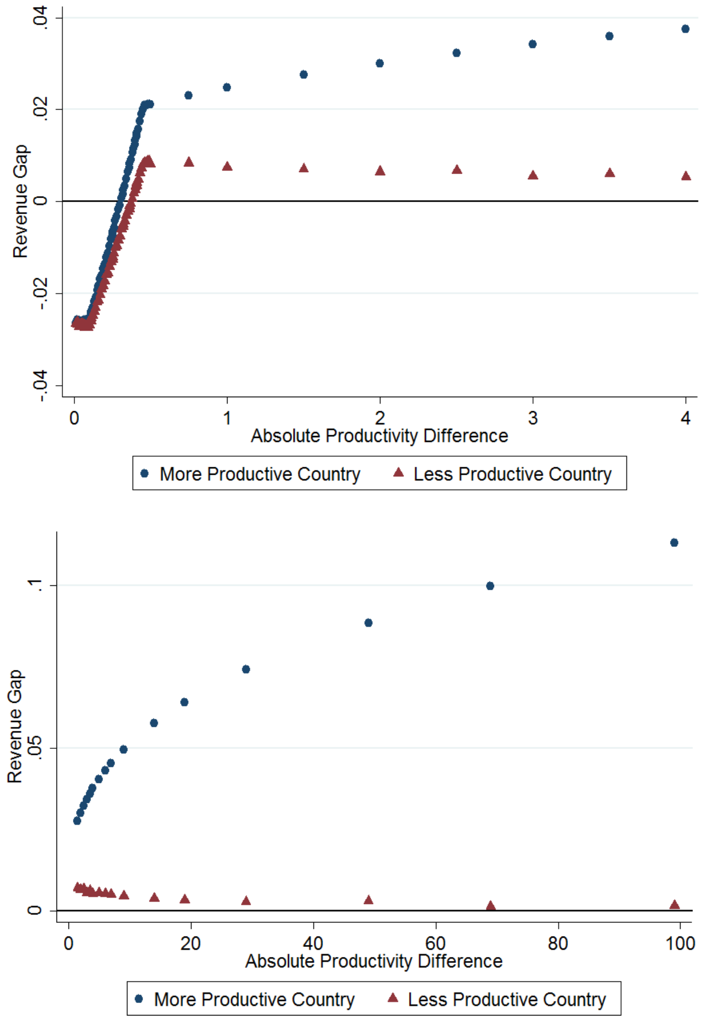


Figure 2.4.6: Revenue Gap and Productivity Asymmetry II

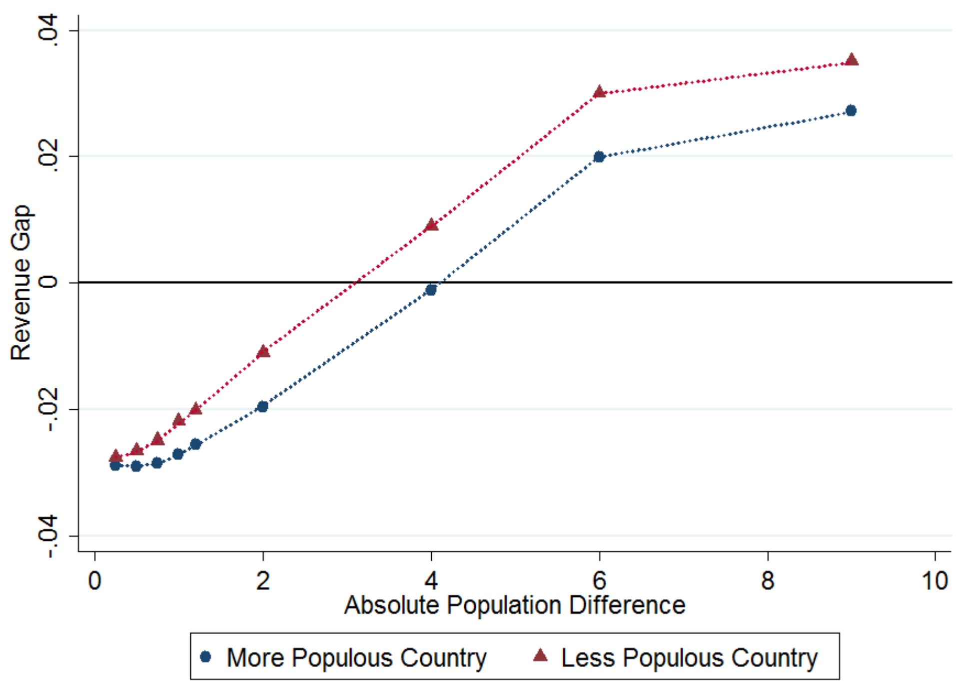


Notice that the revenue gap for both countries is increasing with asymmetry. Initially, the gap is negative which implies that the non-preferential regime is revenue-dominant. When asymmetry rises to 30% and 37%, the preferential and non-preferential regimes generate the same equilibrium revenues for the more and less productive countries, respectively. Above 30%, the more productive country now receives greater revenue under the preferential regime.

Above 37%, both countries find this to be true.[55]

Figure 2.4.6 graph this relationship for higher asymmetries. Notice that the revenue gap for the more productive country is increasing but at a decreasing rate. For the less productive country, the revenue gap eventually starts to decrease but always remains above zero. As moving costs increase, so does the disparity between the two countries. In the case of $c = 0.2$, the more (less) productive country now reaches a revenue gap of zero when productivity asymmetry is at 26% (45%). When including deductibility, increasing α generally decreases the revenue gap except when moving costs are high.

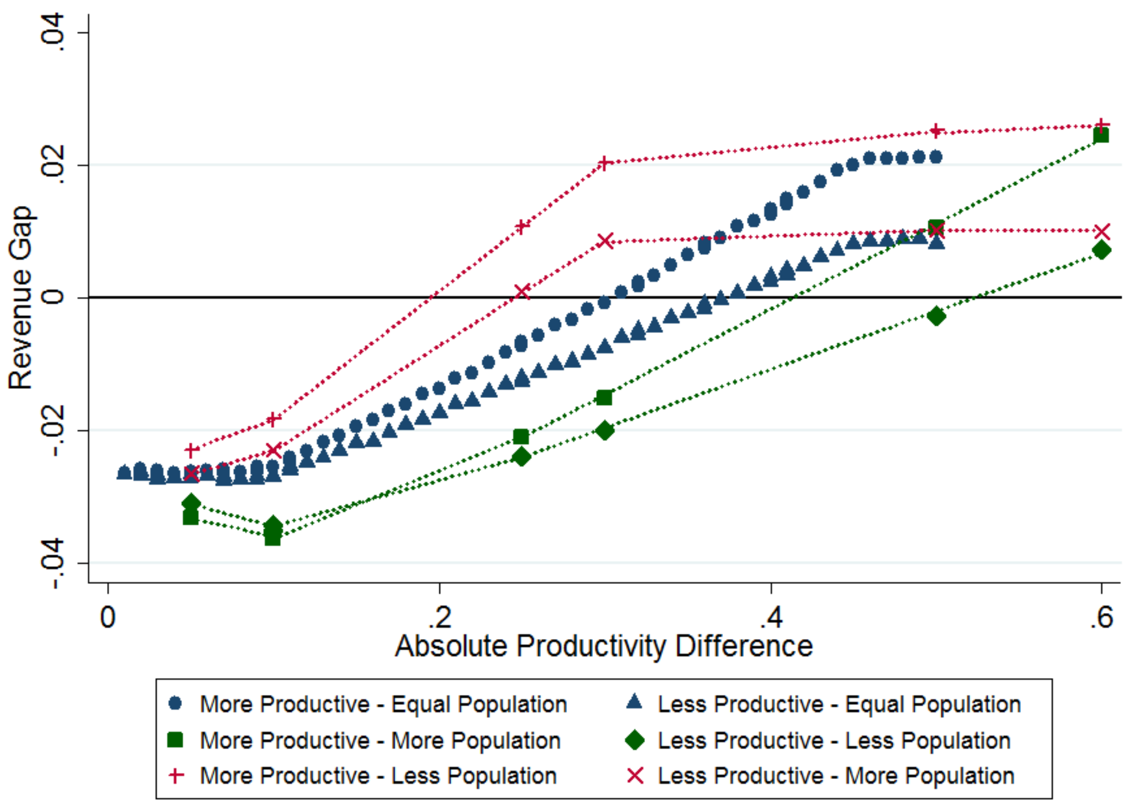
Figure 2.4.7: Revenue Gap and Population Asymmetry



[55]When only one country generates more revenue under the global preferential regime, it is almost always the case that the sum of both their revenue gaps is positive. Because global revenues are higher under the preferential regime, this implies that a revenue transfer exists that would make both countries better off under the preferential regime.

Figure 2.4.7 graphs the analogous relationship if we consider two equally productive countries but with aysmmetric populations. Notice that population asymmetry has a similar albeit far weaker effect. As with productivity, the relationship initially follows a positive, close to linear trend. The less populous country, however, always has a higher revenue gap than the more populous country. As such, the less populous country reaches a revenue gap of zero when population asymmetry is approximately 300% versus the 400% needed for the more populous country. At higher asymmetries, the revenue gaps for both countries fall back towards zero.

Figure 2.4.8: Revenue Gap and Productivity-Population Asymmetry



If we allow for both types of asymmetry simultaneously, the positive effect of productivity asymmetry on the revenue gap is strengthened when the more productive country is less populous. Conversely, greater population in the more productive country dampens this effect. Figure 2.4.8 takes the productivity asymmetry plot in Figure 2.4.5 and overlays two additional plots. The higher pair of lines corresponds to the revenue gaps when asymmetries go in opposite directions. The lower pair of lines corresponds to the revenue gap when asymmetries are aligned.

As an example, consider the case of of $(1, 2, 2, 1)$, $c = 0.1$, and $\alpha = 0$. Country A is less productive but larger in population than country B. Under the preferential regime, A and B receive revenues of $(R_A, R_B) = (0.409, 1.582)$. Under the non-preferential regime, the two countries receive revenues of $(R_A, R_B) = (0.399, 1.55)$ which results in revenue gaps of 0.01 and 0.032 for A and B, respectively. Expressed in percentages, this amounts to a 2.47% and 2% revenue improvement over the non-preferential revenues. The resulting loss gap is equal to -0.0095 which represents a 12.3% difference in TNP loss.

Table 2.21: Preferential Equilibria - $(1, 2, 2, 1)$ $\alpha = 0$

$c =$	t_A	T_A	t_B	T_B	revenue$_A$	revenue$_B$	TNP Loss
0	0.5	0.5	0.673	0.673	0.38226	1.62563	0.06997
0.1	0.575	0.427	0.703	0.641	0.40871	1.58206	0.06785
0.2	0.65	0.36	0.732	0.607	0.43666	1.53656	0.06882
0.3	0.725	0.301	0.763	0.573	0.46888	1.49546	0.07145

Why is the preferential regime only revenue-dominant in these particular cases? When profits are non-correlated, countries have location specific advantages/rents because some firms are inherently more productive in one country. Assuming that Country A is the more productive country, a significant portion of its original firms would generate more profits at home than they would abroad. Therefore, these firms are strongly incentivized to remain in A. Likewise, a significant portion of B's original firms also generate more profits in A. These firms are strongly incentivized to move. Greater productivity asymmetry increases the proportion of such firms. More importantly, both Country A and Country B face fewer marginal firms. Because most of the firms are very likely to locate in A, this reduces the proportion of elastic, tax-sensitive "swing" firms located near the 45° line. This also reduces the benefits of tax competition for Country A. By lowering its rates, it can gain/keep relatively few firms at the cost of losing much greater revenues from its already large contingent of domestic and foreign firms. For Country B, it has a relatively small number of strongly incentivized firms but A's dominance allows it to set higher rates and extract greater revenues.

To see why the preferential regime is better in this setting, consider the benefit and cost of switching regimes. There are two revenue effects when countries switch from non-preferential to preferential taxation. Due to the fact that preferential taxation allows each country to tailor revenue maximizing tax rates for each base, this grants the preferential regime greater revenue generating abilities. At the same time, the strategic interaction aspect, i.e., tax competition, is indeed worse with two tax rates. The sum of these two effects, one positive and one negative, describes the sign of the revenue gap. When countries are symmetric, the negative tax competition effect dominates and makes the total effect/revenue gap negative. For sufficiently asymmetric productivities, the dominance of the more productive country mitigates the negative tax competition effect. Therefore, the total effect/revenue gap becomes positive.

This is more pronounced in the case of a smaller, more productive Country A and a larger, less productive Country B. Recall that Country A is heavily favored and really does not have to worry about tax competition. Therefore, Country A will do better under a preferential equilibrium. For Country B, it will have relatively few remaining firms regardless of the regime. Due to the moving cost and its larger population, B can generate greater revenues by focusing on its domestic base as opposed to using a uniform tax rate to try and mediate between both bases. Thus, the preferential regime is better for Country B as well. In the opposite scenario, a smaller, less productive Country B now must focus on its foreign base since Country A has greater population and thus more firms to attract. Whereas the moving cost made it easier for B to keep its domestic firms in the previous scenario, the moving cost now makes it harder for B to attract foreign firms. Therefore, we need greater asymmetry before Country B's revenue gap becomes positive.[56]

We can also understand this result by looking at the effect of productivity asymmetry on the tax rates. From Section 2, we know that $t \geq \tau \geq T$ is generally true. With higher productivity asymmetry, larger rents allow A to set higher domestic rates as its domestic firms are less likely to exit. It can also set higher foreign rates because foreign firms are more likely to enter. This in turn allows B to increase its own rates. From the perspective of A, greater productivity asymmetry increases its equilibrium domestic and foreign tax rates by a factor of 0.128 and 0.161. The same increase only generates a 0.091 increase in the uniform rate.[57] The preferential regime is therefore able to sustain relatively higher rates than the non-preferential regime which in turn generates greater equilibrium revenues. For Country B, greater asymmetry decreases its uniform rate more than its preferential rates. Thus, an

[56]Assuming that pure strategy equilibria existed, this result suggests that higher profit correlation should theoretically go against the preferential regime because there is less separation between firms. Niu [2013a], Chapter 3, finds that higher profit correlations do in fact decrease the revenue gap. This also suggests that greater ease in profit shifting works against preferential regimes.

[57]These numbers come from a regression of the equilibrium tax rates on productivity and population asymmetry, the moving costs, and moving cost-deductibility interactions. The full set of coefficients is not shown.

increase in productivity asymmetry increases the revenue gap for both countries.

To better quantify the effects of changing asymmetry and moving costs on the loss gap, Table 2.22 presents regression results on the revenue gap in levels and in percentages. Because the observations in the sample are simulated data points, there is no randomness in the regressions.

2.4.4 Less Distortions Under the Preferential Regime

Recall that the loss gap is defined as the difference in TNP loss between the preferential and non-preferential regimes. A positive (negative) loss gap therefore implies that the non-preferential (preferential) regime is less distortionary. As previously discussed, the two regimes generate equal TNP losses when moving costs are zero. When moving costs are positive and profits are non-correlated, the major determinant of the loss gap is the level of moving cost deductibility.

Proposition 4. *The preferential regime is less distortionary, except when moving cost deductibility is high.*

For $\alpha = \{0, 0.25\}$, the loss gap is negative, in all cases of productivity and population asymmetry, and positive moving costs, i.e., the preferential regime is less distortionary at low levels of deductibility. For $\alpha = 0.5$, a small number of cases now exist where the non-preferential regime is less distortionary. Only for $\alpha = \{0.75, 1\}$ do we see a significant shift towards the non-preferential. This stems from the fact that the preferential regime discourages firm movement to a lesser extent than the non-preferential regime. Recall that domestic rates tend to be higher than the uniform rates which tend to be higher than the foreign rates. The combination of higher domestic rates and lower foreign rates creates greater tax-incentives for firms to move under the preferential regime. When deductibility is very high, this causes rates to increase but not enough to prevent too much movement.

Table 2.23 presents analogous regression results on the loss gap in levels and in percentages. As we can see, both higher population and productivity asymmetry are negatively (positively) correlated with the loss gap for the more (less) productive country. The magnitude of the productivity effect, however, is much smaller. In terms of moving costs, higher c decreases the loss gap. This is caused by the sharp decrease in firm movement under the non-preferential regime. When we factor in deductibility, larger α is positively correlated with the gap. Moving costs and deductibility therefore have opposite effects, with moving costs dominating at lower levels of c and deductibility dominating at higher levels.

2.5 Conclusion and Policy Implications

This paper is the first of its kind to look at a more sophisticated and less restrictive model that jointly incorporates population and productivity asymmetry, moving costs, cost deductibility, and non-correlated profits in a heterogeneous firm setting. Due to analytical intractability, the literature lacked a comprehensive model that allowed for these considerations. As such, this paper sheds light on the preferential/non-preferential debate by analyzing a numerical model. Compared to previous papers, the results from this paper are more in line with Janeba and Smart [2003] as opposed to Keen [2001] or Bucovetsky and Haufler [2007] in the sense that the revenue-dominance of a given regime is case dependent and not universal.

Unlike Janeba and Smart [2003], however, this paper finds a much greater number of cases under which the non-preferential regime is not revenue-dominant. When moving costs are zero, the two regimes are nearly identical which implies that regime choice does not matter under perfect mobility. When profits are non-correlated and productivity is sufficiently asymmetric, the preferential regime is now revenue-dominant. This result is strengthened if the more productive country is also less populous. In regards to location choice distortions, the preferential regime skewed firms' decisions to a lesser extent for $\alpha < 0.75$.

These findings have significant policy implications. Take the example of Canada and Spain. Both countries have approximately 2.5 million registered firms. As a rough measure of productivity, both also generate approximately $43 of GDP per labor hour.[58] If we believe that countries such as Canada and Spain are likely to have low levels of productivity and population asymmetry, then the model suggests that greater revenues could be generated if they had non-preferential regimes, i.e., Canada and Spain would not offer preferential tax rates to Canadian or Spanish firms. Conversely, consider the pairing of Belgium and Spain. Belgium has only 350,000 registered firms but generates $54 of GDP per labor hour. If we believe that these two countries are likely to have high levels of productivity asymmetry with opposing population asymmetry, then preferential regimes may result in higher revenues, i.e., Belgium and Spain could offer preferential rates to Belgian and Spanish firms.

Overall, the model suggests that tax regimes should be instituted on a country-by-country basis. Instead of completely banning preferential policies, countries should have specific, bilateral tax regime agreements. This runs counter to the opinions of the EU and OECD. Finally, the model also suggests that differences between regimes are shrinking as moving costs approach zero. As globalization continues and inputs/firms become increasingly mobile, the potential revenue and distortion costs of implementing the "incorrect" regime are also decreasing. However, the ever growing fiscal needs of governments ensure that this regime-revenue debate will remain a pressing and salient issue for the foreseeable future.

[58] Labor productivity numbers, expressed in US dollars, come from the 2010 OECD Factbook. Firm numbers come from the 2005 World Bank Development Indicators.

Table 2.22: Revenue Gap in Levels and Percentages

(Prod Diff ? 0)	Level (>)	Level (<)	Level (=)	Percent (>)	Percent (<)	Percent (=)
Prod Diff	0.074***	-0.041***		0.069***	-0.104***	
	(0.006)	(0.004)		(0.008)	(0.010)	
Pop Diff	-0.020***	0.014***	-0.008***	-0.024***	0.036***	-0.004**
	(0.003)	(0.002)	(0.001)	(0.003)	(0.004)	(0.002)
c = 0.1	-0.028***	-0.023***	-0.027***	-0.027***	-0.055***	-0.062***
	(0.006)	(0.004)	(0.003)	(0.008)	(0.010)	(0.006)
c = 0.2	0.008	-0.018***		0.012	-0.042***	
	(0.006)	(0.004)		(0.008)	(0.010)	
c = 0.3	-0.006	-0.044***	-0.144***	-0.009	-0.085***	-0.268***
	(0.007)	(0.005)	(0.004)	(0.009)	(0.012)	(0.007)
c = 0.1 a = 0.25	-0.005	-0.002	-0.005**	-0.005	-0.006	-0.011**
	(0.006)	(0.004)	(0.003)	(0.008)	(0.010)	(0.005)
c = 0.1 a = 0.5	-0.011	-0.004	-0.008***	-0.010	-0.012	-0.017***
	(0.006)	(0.004)	(0.003)	(0.008)	(0.011)	(0.005)
c = 0.1 a = 0.75	-0.014**	-0.005	-0.008***	-0.012	-0.015	-0.018***
	(0.006)	(0.004)	(0.003)	(0.008)	(0.011)	(0.005)
c = 0.1 a = 1	-0.016**	-0.005	0.001	-0.014	-0.016	-0.001
	(0.007)	(0.004)	(0.003)	(0.008)	(0.011)	(0.005)
c = 0.2 a = 0.25	-0.022***	-0.011**	-0.029***	-0.024***	-0.031***	-0.063***
	(0.007)	(0.004)	(0.003)	(0.008)	(0.011)	(0.006)
c = 0.2 a = 0.5	-0.041***	-0.020***	-0.051***	-0.046***	-0.055***	-0.109***
	(0.007)	(0.004)	(0.003)	(0.008)	(0.011)	(0.006)
c = 0.2 a = 0.75	-0.051***	-0.023***	-0.065***	-0.056***	-0.064***	-0.137***
	(0.006)	(0.004)	(0.003)	(0.008)	(0.011)	(0.006)
c = 0.2 a = 1	-0.052***	-0.020***	-0.053***	-0.057***	-0.059***	-0.115***
	(0.007)	(0.004)	(0.003)	(0.008)	(0.011)	(0.006)
c = 0.3 a = 0.25	0.050***	0.026***	0.001	0.063***	0.044***	-0.018*
	(0.008)	(0.005)	(0.005)	(0.010)	(0.013)	(0.010)
c = 0.3 a = 0.5	-0.005	-0.000	0.120***	0.001	-0.028**	0.230***
	(0.008)	(0.005)	(0.006)	(0.010)	(0.012)	(0.011)
c = 0.3 a = 0.75	-0.039***	-0.017***	0.075***	-0.039***	-0.070***	0.126***
	(0.008)	(0.005)	(0.005)	(0.010)	(0.012)	(0.009)
c = 0.3 a = 1	-0.044***	-0.014***	0.066***	-0.046***	-0.066***	0.109***
	(0.008)	(0.005)	(0.004)	(0.010)	(0.012)	(0.007)

The dependent variables are the revenue gaps in levels and percentages of the non-preferential revenue. The baseline case is zero moving costs and zero deductibility. Columns are separated for more, less, and equally productive countries.

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

Table 2.23: Loss Gap in Levels and Percentages

(Prod Diff ? 0)	Level (>)	Level (<)	Level (=)	Percent (>)	Percent (<)	Percent (=)
Prod Diff	-0.001	0.001		-0.771**	0.771**	
	(0.003)	(0.003)		(0.385)	(0.385)	
Pop Diff	-0.010***	0.010***	0.000	-0.343**	0.343**	0.000
	(0.001)	(0.001)	(0.000)	(0.167)	(0.167)	(2.442)
c = 0.1	-0.008***	-0.008***	0.029***	0.212	0.212	-0.013
	(0.003)	(0.003)	(0.001)	(0.373)	(0.373)	(7.060)
c = 0.2	-0.040***	-0.040***		0.032	0.032	
	(0.003)	(0.003)		(0.379)	(0.379)	
c = 0.3	-0.046***	-0.046***	-0.046***	0.010	0.010	-0.068
	(0.004)	(0.004)	(0.001)	(0.452)	(0.452)	(8.930)
c = 0.1 a = 0.25	0.005	0.005	0.006***	0.104	0.104	0.138
	(0.003)	(0.003)	(0.001)	(0.405)	(0.405)	(6.061)
c = 0.1 a = 0.5	0.008**	0.008**	0.010***	0.225	0.225	0.504
	(0.003)	(0.003)	(0.001)	(0.409)	(0.409)	(6.235)
c = 0.1 a = 0.75	0.010***	0.010***	0.013***	0.326	0.326	2.322
	(0.003)	(0.003)	(0.001)	(0.409)	(0.409)	(6.235)
c = 0.1 a = 1	0.010***	0.010***	0.013***	0.381	0.381	53.389***
	(0.003)	(0.003)	(0.001)	(0.410)	(0.410)	(6.235)
c = 0.2 a = 0.25	0.017***	0.017***	0.012***	0.085	0.085	0.023
	(0.004)	(0.004)	(0.001)	(0.411)	(0.411)	(7.337)
c = 0.2 a = 0.5	0.031***	0.031***	0.029***	0.241	0.241	0.208
	(0.003)	(0.003)	(0.001)	(0.410)	(0.410)	(7.862)
c = 0.2 a = 0.75	0.042***	0.042***	0.044***	0.577	0.577	1.221
	(0.003)	(0.003)	(0.001)	(0.408)	(0.408)	(7.669)
c = 0.2 a = 1	0.046***	0.046***	0.050***	1.043**	1.043**	48.647***
	(0.003)	(0.003)	(0.001)	(0.410)	(0.410)	(7.528)
c = 0.3 a = 0.25	0.012***	0.012***	0.000	0.146	0.146	-0.043
	(0.004)	(0.004)	(0.002)	(0.509)	(0.509)	(11.981)
c = 0.3 a = 0.5	0.020***	0.020***	0.076***	0.178	0.178	0.448
	(0.004)	(0.004)	(0.002)	(0.479)	(0.479)	(13.906)
c = 0.3 a = 0.75	0.045***	0.045***	0.089***	0.534	0.534	0.704
	(0.004)	(0.004)	(0.002)	(0.476)	(0.476)	(11.028)
c = 0.3 a = 1	0.058***	0.058***	0.106***	2.759***	2.759***	47.722***
	(0.004)	(0.004)	(0.001)	(0.474)	(0.474)	(8.472)

The dependent variables are the loss gaps in levels and percentages of the non-preferential TNP loss. The baseline case is zero moving costs and zero deductibility. In ten cases, the non-preferential TNP loss is zero. The percentage loss gap is coded as 100 (the max is 86.88). Columns are separated for more, less, and equally productive countries.

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

CHAPTER III

Preferential Corporate Taxation and Profit Correlation Under a Bivariate Pareto Distribution

3.1 Introduction

Preferential taxation refers to the usage of base discriminatory tax rates as opposed to non-preferential taxation which imposes a single, uniform tax rate on all bases. This discrimination can be sector-specific as in the case of Ireland which gave its manufacturing sectors, amongst others, a 10% versus the standard 32% corporate income tax rate in the 1980's and 1990's. The true purpose of Ireland's preferential rates, however, was to attract a greater number of foreign firms and capital, as the targeted sectors had significant foreign involvement and investment. Therefore, this sector-based discrimination was, in actuality, thinly veiled source-based discrimination to increase the foreign tax base.

The success of Ireland and other countries led to an increased usage of preferential taxation. Policymakers in the EU and OECD, however, feared that preferential treatment of more mobile bases would result in competition-driven rate undercutting and an eventual race to the bottom. They therefore published a number of reports in the early 2000's identifying and reprimanding countries using such practices. Since then, many of the offending preferential tax policies have been either abolished or amended. Ireland, for example, abandoned its sector-specific rate and instead imposed a uniform, 12% corporate tax rate for all firms. However, source-specific preferential taxation still exists most commonly in the form of FDI-friendly special economic zones in countries such as China and India. These SEZs offer amenities such as lower statutory rates, special deductions, and temporary tax holidays for qualifying foreign firms.

Tax competition is not a new concept in the economic literature. Oates [1972] discussed inefficiencies in the context of local governments competing over mobile capital. The notion that inter-jurisdictional competition could lead to inefficiently low taxation and misallocations contrasted greatly with the prevailing Tiebout [1956] hypothesis. Emphasis on the equilibrium revenue-dominance of preferential versus non-preferential regimes, however, is fairly recent. Intuitively, all countries have an incentive to *unilaterally* adopt a preferential regime with the rationale being that more tax instruments, i.e., flexibility, cannot result in strictly less revenue. However, allowing all countries to adopt preferential policies may result in lower levels of equilibrium revenues for all, similar to the Prisoner's Dilemma result. The choice, therefore, centers on whether allowing *all* countries to adopt preferential regimes generates more or less per-country equilibrium revenues than the counterfactual equilibrium where *all* countries have non-preferential regimes.

The two primary voices/papers in this debate are Keen and Janeba. While papers such as Janeba and Peters [1999], which looked at the choice of regimes from a sequential setting, tended to support non-preferential taxation for the most part, Keen [2001] presented a simple two (symmetric) country, two base model that pointed towards the revenue-dominance of the preferential regime. The rationale behind his result was that having a separate domestic versus foreign tax rate shielded the more immobile domestic base from the negative effects of tax competition. Under non-preferential taxation, a uniform tax rate negatively effects both bases. The preferential regime, however, focuses competition onto the more mobile base while allowing the less mobile base to be heavily taxed. This presented a shock to the literature which prompted other papers to build on or re-evaluate Keen's result. The most prominent paper against the Keen [2001] result was Janeba and Smart [2003]. Under a more generalized model which allowed for perfectly immobile firms and aggregate base effects, i.e., higher tax rates could not only push capital to a lower jurisdiction but also into other non-taxed sectors, the authors found that Keen's model and result could be represented as a specific case under their larger model. In most cases, however, Janeba and Smart [2003] found that banning preferential taxation resulted in higher equilibrium revenues. Subsequent works by others such as Haufler, Haupt, and Wilson have tried to amend these two basic models.[59]

This paper builds on my previous Niu [2013b], Chapter 2, model which developed a richer, more realistic, and less constrained setup that incorporated firm heterogeneity, country-level productivity and population asymmetries, as well as moving costs and cost deductibility. Central to this paper, it also introduced the concept of cross-country profit correlation. The standard assumption in the literature is that profits, returns, etc. are independent of the country/location choice excluding differences in tax rates and moving costs, i.e., perfect one-to-one profit correlation. Niu [2013b], however, analyzed the case of perfect non-correlation where a firm's profits in different countries are independent and uniformly distributed. In

[59]See Genschel and Schwarz [2011] for a review of the literature.

doing so, the paper found that the preferential regime was revenue-dominant in a large number of cases where profits were perfectly non-correlated.

Niu [2013b] was only able to model the two polar cases of perfect correlation and perfect non-correlation. The question remains as to the revenue-regime relationship in between these two extremes. This paper re-examines these results by focusing on the equilibrium revenues generated under varying levels of imperfect profit correlation. I do so by constructing and simulating a numerical, copula-based model that incorporates three aspects from the international literature. (1) Firms' profits follow Zipf's Law such that the majority of firms make very little profit with only a small fraction making large profits.[60] (2) Across the OECD countries, these profits/firms follows the same Pareto distribution in all countries.[61] (3) Because the distributions are the same, differences in profitability/productivity across countries arise due to variation in draws as well as differences in the population of firms, i.e., the number of draws. I therefore construct a two country model where firms' profits are drawn from two identical bivariate Pareto distributions. To allow for varying levels of profit correlation, I generate the bivariate Pareto distributions using a heavy-right tail (HRT) copula. This copula construction allows me to specifically adjust the correlation between cross-country profits without changing the underlying Pareto marginals.

The simulation results suggest that the effects of imperfect profit correlation on the equilibrium revenue differences between the two regimes are highly non-linear. The relationship is approximated using a cubic polynomial which indicates that the direction of marginal effects from small changes in correlation depend on the current correlation. Additionally, the effect of correlation is highly dependent on the nature of the correlation - own country

[60] Axtell [2001] and Fujiwara et al. [2004] both find that firm size, as measured using a number of different proxies including receipts, follows a power law which is Zipf in many instances.

[61] See the empirical estimates from di Giovanni et al. [2011].

versus other country correlation. Overall, however, average increases in the level of profit correlation tend to decrease the per-country equilibrium revenues generated under the preferential regime relative to the non-preferential regime. Lastly, results from the simulation show the asymptotic versus small sample disparities between Niu [2013b] and this paper. While Niu [2013b] analyzed a model with a continuum of firms, this paper simulates a finite sample scenario. Therefore, results from the asymptotic or large sample model do not fully carry-over to this smaller, stochastic model. Most notably, I no longer find that differences between the regimes are zero under perfect mobility.

The remainder of the paper proceeds as follows. Section 2 describes the model, including its relationship to Keen [2001] and Niu [2013b]. Section 3 discusses the simulation setup, parameter values and key variables, and provides relevant summary statistics. Section 4 presents the results of the simulation and Section 5 concludes.

3.2 Model

Following the basic model in Keen [2001], assume that there exist two countries, A and B, and two tax bases with varying levels of mobility. Under a territorial system, countries are only able to tax the fraction of each base located within their borders. Notation-wise, we differentiate between rates on the two bases using t_j and T_j for $j = \{A, B\}$. In this one-period, simultaneous game, the two countries set tax rates on the two bases so as to maximize revenue. After observing the rates, the bases react and move accordingly. The revenue objective is both a simplifying and realistic assumption. Maximizing revenue requires fewer assumptions over social welfare, labor distortions, and potential general equilibrium considerations such as agglomeration. Additionally, the revenue maximization objective is arguably the correct and true objective of policymakers.

3.2.1 Firms' Problem

In the same vein as Niu [2013b], I augment the simple Keen model to include more sophisticated and detailed considerations. First, assume that the two bases are comprised of stocks of atomistic, profit-maximizing firms. Originally, i.e., prior to firm movement, each country is endowed with a population of firms. In a heterogeneous firms setting, the profit that firm i makes in country j, π_{ij}, is assumed to follow a Pareto distribution. This is true for firms originally located in either Country A or B. Each firm is therefore characterized by its *profit pair*, (π_A, π_B), which denotes the profit that it would make if it were located in each corresponding country. In regards to the timing of the model, assume that firms receive their profit pair draws prior to deciding on their location choices and after countries' tax rates have been set. This implies that firms are operating under full information. They know exactly how much gross profit they would generate and how much net (of tax) profit they would retain before deciding whether to stay in their country of origin or relocate abroad. As is standard, firms act to maximize net profits. For a firm originally located in Country A, it would receive a net profit of $(1 - t_A)\pi_A$ where t_A refers to the corporate profit tax rate on *domestic* firms in A. If the firm moved to Country B, it would receive a net profit of $(1 - T_B)\pi_B - c$ where T_B refers to the corporate profit tax rate on *foreign* firms in B. c is a fixed moving cost and is constant across countries and firms.[62]

As previously mentioned, both π_A and π_B follow identical Pareto distributions such that the potential profit pairs in each country are drawn from identical *bivariate* Pareto distributions, $f(\pi_A, \pi_B)$. The literature typically assumes that cross-country profits follow a perfect,

[62]In the context of start-up, the question arises as to why $c > 0$ is reasonable if the firm has yet to establish a location. Given the example of an entrepreneur who has lived in the United States, choosing to start in Germany could still involve a positive moving cost which we can interpret as the dollar hassle cost of having to learn and/or adjust to new laws, best practices, etc.

one-to-one profit correlation, i.e., $\pi_{iA} = \pi_{iB}$ for all firms i. Distributionally, this means that $f(\pi_A, \pi_B) = 0$ in all cases where $\pi_A \neq \pi_B$. Niu [2013b] also looked at the other extreme case of perfect non-correlation such that π_A and π_B were independent of each other. The independence implied that the profitability of a given firm in one country provides no information regarding its profitability in the other country.

This paper/model analyzes levels of imperfect correlation, as it would be plausible to assume that a highly profitable firm in Country A would likely be highly profitable as well in Country B, at least to some extent. To do so, I construct the bivariate Pareto distribution using a heavy-right tail (HRT) copula.[63] Derived from Sklar's Theorem, a copula is simply a functional form that can model, and more importantly, vary the degree of dependency between multiple, univariate random variables. Written in terms of the marginal distribution functions of any set of univariate variables, a copula is able to generate a cumulative multivariate distribution function denoted by $C(u_1, u_2, ...u_i)$.[64] The cumulative density function of the HRT copula is given by:

$$F(\pi_A, \pi_B) \equiv C\left(u_A(\pi_A), u_B(\pi_B)\right) = u_A + u_B - 1 + \left[(1-u_A)^{-\frac{1}{\rho}} + (1-u_B)^{-\frac{1}{\rho}} - 1\right]^{-\rho} \quad (3.2.1)$$

The u terms represent the marginal univariate profit distributions in Countries A and B. Because I have set these to be univariate Pareto distributions following the international literature, let $u_j = 1 - \pi_j^{-\theta}$ such that $\pi_j \in [1, \infty)$ for $j = \{A, B\}$. In each distribution, the θ term, bounded from below at one, determines how quickly the density tails off for higher profits. As θ increases, the tails shrink and the probability of receiving a high profit

[63] The HRT copula is equivalent to a flipped-Clayton copula.

[64] The application of copulas is common in finance as well as in actuarial work to model and predict the likelihood of multiple events occurring together, e.g., dips in stock prices or the occurrence of multiple accidents.

draw decreases. θ is set to 1.05 in this model following empirical estimates from previous literature. The most important term in this HRT copula is the ρ term. It parameterizes the level of correlation between the two univariate profit distributions. As $\rho \to 0$, the distribution approaches perfect correlation. As $\rho \to \infty$, the bracketed term asymptotes towards one. This results in perfect non-correlation as we now have two separable (independent) univariate distributions.

For this paper, the usage of a copula is crucial for one main reason. It allows me to separately and independently control the correlation between π_A and π_B. Consider the density function for the Type I distribution from Mardia [1962], which is also a bivariate Pareto distribution.

$$f\left(\pi_A, \pi_B | a, b, p\right) = \frac{p\left(p+1\right)\left(ab\right)^{p+1}}{\left(b\pi_A + a\pi_B - ab\right)^{p+2}} \tag{3.2.2}$$

Analogously, the p term in (3.2.2) controls the correlation between π_A and π_B. However, changing the p term also affects the means of the two random variables as seen below.

$$\mathrm{E}\left(\pi_j\right) = \frac{ap}{p-1} \tag{3.2.3}$$

Because the goal is to compare differences in equilibrium revenues between the two regimes at varying levels of correlation, we need to keep other factors, such as the mean, constant. The HRT copula allows us to vary the correlation while holding the underlying distributions fixed. This facilitates the analysis of the effect of imperfect correlation and only correlation on the regime choice.

Given the profit pairs, firms are profit maximizing, i.e., the location decision is based on the net-of-moving cost, net-of-tax profits afforded in each country. Thus, a firm originally in

Country A will move to Country B if the following inequality holds.

$$\pi_B - T_B\pi_B - c > \pi_A - t_A\pi_A \tag{3.2.4}$$

$$\pi_B > \frac{(1-t_A)\,\pi_A + c}{1-T_B} \tag{3.2.5}$$

3.2.2 Governments' Problem

As previously mentioned, each country's government has access to its respective domestic and foreign tax rates, t_j and T_j. Under the preferential regime, t_j and T_j are allowed to be different. Under the non-preferential regime, the two tax rates must be equal. Define this equal, uniform tax rate as $\tau_j \equiv t_j = T_j$. Given these tax rates and constraints, the objective of Country A is to maximize the sum of taxes collected from its domestic and foreign firms/bases. Likewise, Country B has analogous objective functions.

$$\max_{t_A, T_A} R_A = R_A^{\text{domestic}}\,(t_A, T_B, c, \rho) + R_A^{\text{foreign}}\,(T_A, t_B, c, \rho) \tag{3.2.6}$$

$$\max_{\tau_A} R_A = R_A^{\text{domestic}}\,(\tau_A, \tau_B, c, \rho) + R_A^{\text{foreign}}\,(\tau_A, \tau_B, c, \rho) \tag{3.2.7}$$

Countries A and B set their tax rates prior to the realization of firms' profit pairs. Therefore, tax rates are determined based solely on the distributional characteristics of firms. This prevents countries from setting firm specific rates. In this single period setup, I assume that tax rates are set simultaneously so comparisons between regimes will look at the relative levels of pure strategy Nash Equilibrium revenues for each country. Below, I detail the simulation parameters as well as the computational methodology.

3.3 Simulation Setup

Recall that the central research question of this paper pertains to the effect of profit correlation on equilibrium revenues under the two regimes. What we are interested in is the comparison of the per-country equilibrium revenues when all countries are utilizing a preferential regime versus when all countries are restricted to a non-preferential regime. To this end, I simulated a numerical version of the model using MATLAB. The level of profit correlation, ρ, ranged between 0.05 and 20. Recall that a smaller ρ implies that the differences in firms' cross-country profits are smaller, i.e., are more correlated. The marginal correlation effect of increasing ρ decreases quickly at values of ρ greater than five. I therefore placed more emphasis on smaller values of ρ in the cases considered.[65] In regards to moving costs, I varied c between $c = \{0, 1, 2, ..., 10\}$. c is capped at 10 due to the fact that higher moving costs result in a significant majority of relatively immobile firms. Equilibrium tax rates in these higher cases approached or hit 100%.

The simulations also accounted for differences in population and productivity. In the symmetric case, both countries started with a stock of $N_A = N_B = 10,000$ firms/draws from the distribution. I also simulated the asymmetric cases where N_A was allowed to vary between 10,000 and 20,000 firms in increments of 2,500. The population difference between the two countries (in thousands of firms) therefore ranged between -10 and 10. Because more draws increases the likelihood of getting a single high profit firm, these asymmetric populations are also a proxy for asymmetric productivities.

For each combination of ρ and population, I drew N_A and N_B random samples from the HRT constructed bivariate Pareto distribution. These draws/profit pairs became the original stocks of firms in each country within the simulated two country world. Due to the fact that

[65] $\rho = \{0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.3, 1.5, 2, 3, 4, 5, 8, 10, 15, 20\}$

these draws are random, it is dangerous to draw conclusions solely off of one "observation" as sample parameters may differ significantly from population parameters. I therefore re-drew the populations for each of the above cases under fifty different seed combinations to mitigate small sample effects that may have led to spurious results.[66] The total number of simulated cases is therefore 66,000.

To determine the resulting equilibria, an iterated grid-search approach is utilized. See Appendix G for a more detailed explanation of the grid-search methodology. Given an initial starting tax rate in Country B, Country A pins down its optimal tax response. Tax rates are constrained to be between 0 and 1, i.e., 0% and 100%, in steps or increments of 0.0001, i.e., 0.01% points. The process then reverses and Country B now searches for the optimal rate response to A's previous response. This is repeated until neither country has a beneficial deviation. Because we may be worried about identifying the local as opposed to global optima, I re-ran this iterated grid-search from a number of different starting points. This iterated approach identifies only stable equilibria. Unstable equilibria are ignored as they are potentially unrealistic due to trembling hand vulnerabilities.

I run the iteration process under the preferential regime, where both countries are allowed to set separate domestic and foreign rates, and again under the non-preferential regime when countries are only allowed set country-specific uniform rates. This gives us four values for each simulated case: the two equilibrium revenues for Countries A and B when both are optimizing under the preferential regime, and the two equilibrium revenues for Countries A and B when both are optimizing under the non-preferential regime. In total, I generated 132,000 simulated data points. Table 3.24 provides a summary of the equilibrium rates and revenues under the two regimes.

[66]Thirty to thirty-five observations is the commonly accepted threshold for the application of the law of large numbers.

Table 3.24: Summary Statistics

	MIN	**MAX**	**MEAN**	**STD**
t	4.7363%	100%	72.30804%	20.56718%
T	4.8163%	99.995%	71.85129%	19.45502%
τ	5.6765%	99.86%	71.34218%	20.81503%
Pref Rev	4,656.376	6,118,639	111,278.5	134,890.6
Non-Pref Rev	4,609.917	6,103,578	107,657.6	133,992.1
Percent Corr.	6.14828%	5337.793%	132.9909%	1969.488%
Level Corr.	0.4184916	354.1895	7.630336	8.699364
Kendall's Tau	0.0058533	0.9110343	0.3776516	0.2650142

3.4 Results

3.4.1 Measuring Profit Correlation

To quantify the level of profit correlation, I use three different measures: Kendall's Tau, level deviation, and percent deviation. Table 3.24 provides descriptive statistics on these measures which are discussed below.

Kendall's Tau is a measure of rank correlation, i.e., how close the π_A and π_B are in regards to their relative positions within each univariate distribution. It is therefore bounded between -1 and 1. A Tau of 1 and -1 implies perfect positive and negative rank correlation, respectively. A Tau of zero implies rank independence. For example, if the profits are such that $\pi_A \in \{1,2\}$ and $\pi_B \in \{11,12\}$ then a profit pair of $(1,11)$ and $(2,12)$ would imply a Tau of 1. Because these simulations focus solely on positive dependencies, Kendall's Tau is strictly positive within the sample.

In the context of copulas, Kendall's Tau is a copula-specific function of the ρ term. Importantly, it is independent of the marginals, i.e., the u terms from (3.2.1). In the HRT copula, Kendall's Tau is equal to $1/(2\rho+1)$. As ρ ranges between 0.05 and 20 in the simulations, the sample Taus in Table 3.24 line up very closely to the true Taus. Using Spearman's Rho produces similar results. As the marginal, univariate distributions are both Pareto distributions with an identical lower bound set at $\pi = 1$, rank correlation should produce similar results to level measures of correlation. However, this is dependent on having a large sample size. Therefore, two alternative measures of correlation are constructed for the analysis - level deviation and percent deviation.

As opposed to measuring the rank correlation, the level deviation between profits within each profit pair looks at the magnitude of differences between cross-country profits. I define the level deviation as the sample average of the absolute value of the differences in each profit pair. For a sample of N_j draws in country j, the level correlation would be calculated as:

$$\text{Level Deviation} = \frac{1}{N_j}\sum_{i=1}^{N_j} |\pi_{iA} - \pi_{iB}| \tag{3.4.1}$$

Recall from the previous example that the two profit pairs $(1,11)$ and $(2,12)$ have a rank correlation of 1, i.e., we have perfect positive correlation. However, the level deviation is equal to ten because the profit difference is ten in both pairs. A level deviation of zero therefore implies that the cross-country profits for each firm are equal, i.e., perfectly, one-to-one correlated. Higher level deviation values imply that the profits are on average further apart.

Notice that level deviation does not distinguish between positive and negative correlations. Because firms care about profit and not rank, this may be a more appropriate measure, irregardless of sample size.

To account for proportional effects, I also consider percent deviation. This is calculated by averaging the percentage cross-country profit differences as a fraction of π_A or π_B, depending on the country of origin.

$$\text{Percent Deviation} = \frac{1}{N_j} \sum_{i=1}^{N_j} \frac{|\pi_{iA} - \pi_{iB}|}{\pi_{ij}} \tag{3.4.2}$$

Intuitively, this is meant to capture the fact that a profit difference of ten should imply less correlation between π_A and π_B when the baseline profit is 20 versus 20,000. Similar to level deviation, a percent deviation of zero implies that the profits are equal for every pair while higher percent deviation implies less proportional correlation. Overall, notice that with Kendall's Tau, a higher Tau is associated with greater correlation. Conversely, a higher level and percent deviation is associated with less correlation.

In the analysis, the three measures of correlation are further categorized to be own or other correlation to represent profit correlation in one's own country versus the correlation in the other country.

3.4.2 Rates and Revenue

To analyze and summarize the data, OLS regressions are utilized. These regressions are run on samples split up based on whether PopDiff > 0 or PopDiff ≤ 0, i.e., for the more and weakly less populous countries, respectively. There is no uncertainty in the samples aside from the randomness generated by the different seeds. For all regressions, seed dummies,

$seed_i \in \{1, 2, ..., 50\}$, as well as country dummies are included as controls but omitted from the regression results. Likewise, seed-country interactions are also included and omitted.

Table 3.25 presents the regression results on the effects of own and other percent deviation/correlation, population difference, and moving costs on the equilibrium rates. Under the preferential regime, domestic rates increase as both own or other percent deviation increase, i.e., when correlation decreases. A decrease in own correlation generates both greater incentives and dis-incentives for firm movement because $\pi_A - \pi_B$ may become more positive or negative. As long as decreases in correlation do not skew in a specific direction, changes to the likelihood of firm movement should, on average, balance out. Therefore, domestic rates should, at worst, stay relatively constant if not increase under the presence of positive moving costs. A decrease in other correlation acts in a similar fashion.

For more (less) populous countries, the effect of PopDiff is negative (positive). Theory suggests that changes in population should not affect the preferential rates if the changes are proportional across all productivities - Niu [2013b], Wilson and Mongrain [forthcoming]. Because population also affects the range of firm profits in these simulations, the coefficients on PopDiff are very small in magnitude and essentially equal to zero. Moving costs increase domestic rates because they reduce the benefits of moving for all firms. In regards to the uniform tax rate under the non-preferential regime, decreases in correlation, moving costs, and population differences all increase the uniform rate. Likewise, Table 3.26 presents regression results pertaining to the effects of these factors on the equilibrium revenues under the two regimes.

Table 3.25: Effects on the Three Equilibrium Rates

(Pop Diff ? 0)	Domestic (>)	Domestic (<=)	Foreign (>)	Foreign (<=)	Uniform (>)	Uniform (<=)
Own Percent	0.0264***	0.0512***	0.0211***	0.0404***	0.0266***	0.0505***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Oth Percent	0.0476***	0.0323***	0.0454***	0.0271***	0.0508***	0.0322***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pop Diff	-0.0011***	0.0010***	-0.0005**	0.0005***	0.0012***	0.0025***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
C	0.0185***	0.0118***	0.0094***	0.0163***	0.0135***	0.0114***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.5931***	0.6203***	0.5314***	0.4886***	0.5285***	0.5361***
	(0.005)	(0.009)	(0.005)	(0.009)	(0.005)	(0.010)
Observations	52,800	79,200	52,800	79,200	52,800	79,200
R-squared	0.433	0.487	0.405	0.402	0.405	0.425

This table shows regressions results on the three equilibrium rates: the domestic and foreign rates under the preferential regime, and the uniform rate under the non-preferential regime. The regressions are divided between samples where the population difference is greater than zero and less than or equal to zero. Not shown are seed, country, and seed-country interaction terms included in each regression. Using alternate correlation parameters do not substantively change any coefficients. Columns are separated for larger and weakly smaller countries.

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

Table 3.26: Effects on the Two Equilibrium Revenues

(Pop Diff ? 0)	Pref. Rev (>)	Pref. Rev (<=)	Non-Pref. Rev (>)	Non-Pref. Rev (<=)
Own Percent	4,679.0300***	-20,864.6720***	4,283.3778***	-21,980.0830***
	(80.627)	(211.108)	(82.533)	(211.525)
Oth Percent	21,818.8879***	73,914.1774***	21,719.6510***	73,324.9854***
	(100.109)	(178.586)	(102.475)	(178.939)
Pop Diff	5,249.1034***	-4,056.3740***	5,220.0619***	-4,040.1157***
	(54.872)	(95.604)	(56.168)	(95.793)
C	1,395.3293***	283.4013***	952.7234***	-284.5414***
	(48.475)	(97.527)	(49.620)	(97.720)
Constant	40,468.4348***	24,337.1607***	38,920.5619***	26,497.4752***
	(1,167.473)	(5,369.959)	(1,195.058)	(5,380.581)
Observations	52,800	79,200	52,800	79,200
R-squared	0.669	0.729	0.654	0.724

This table shows regressions results on the two equilibrium revenues. The regressions are divided between samples where the population difference is greater than zero and less than or equal to zero. Not shown are seed, country, and seed-country interaction terms. Columns are separated for larger and weakly smaller countries.

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

Table 3.27: Revenue Gap with Indicator Moving Costs

	Rev. Gap
Pop Diff	23.647***
	(7.187)
Own Kendall	-2,216.030***
	(95.891)
C = 1	939.769***
	(119.180)
C = 2	2,074.047***
	(119.180)
C = 3	2,974.695***
	(119.180)
C = 4	3,599.087***
	(119.180)
C = 5	3,984.466***
	(119.180)
C = 6	4,523.552***
	(119.180)
C = 7	4,863.480***
	(119.180)
C = 8	5,002.156***
	(119.180)
C = 9	5,183.517***
	(119.180)
C = 10	5,299.505***
	(119.180)
Observations	132,000
R-squared	0.059

Not shown are seed, country, seed-country interaction terms, and constant.
*** p<0.01, ** p<0.05, * p<0.1

3.4.3 Revenue Gap

Following Niu [2013b] this paper has two primary, dependent variables of interest. The first variable is the revenue gap which I define as the difference in equilibrium revenues generated by the preferential and non-preferential regimes. A positive revenue gap implies that the preferential regime is revenue-dominant while a negative revenue gap implies that the non-preferential regime is revenue-dominant. The second variable is the revenue gap in percentages. It is the revenue gap as a fraction of the non-preferential revenue. Because revenue levels may inherently increase or decrease given changes to other parameters, the percentage revenue gap scales these effects.

As with before, I use OLS regressions to quantify effects on the revenue gap. On the right hand side, the main independent variables of interest are the three correlation measures. Controls used in the regressions include the difference in population, moving costs, seed and country dummies, and seed-country interactions. Quadratic and cubic moving cost terms are included in the regressions to reflect non-linearities as evidenced by Table 3.27. With a moving cost of zero as the baseline, we see that the differential effect of higher moving costs is initially increasing, peaks around $c = 2$, and then diminishes at higher costs. Adding a quartic moving cost term results in insignificant coefficients on both the quartic and cubic terms.

Regressing the revenue gap on own and other correlation suggests that on average, decreasing correlation results in greater revenue. This implies that higher OwnPercent, higher OwnLevel, and lower OwnKendall values make the preferential regime more revenue-dominant. Similar to moving costs, this effect is non-linear. Tables 3.28 and 3.29 present regressions that include splines for the three correlation measures. F-tests on the equality of the spline coefficients allow us to strongly reject the null hypothesis. As is evident, the

Table 3.28: Revenue Gaps with Splines

(Pop Diff ? 0)	Percent (>)	Percent (<=)	Level (>)	Level (<=)	Kendall (>)	Kendall (<=)
Own - 1	-136,444***	-137,173***	-3,299***	-1,525***	45,335***	45,847***
	(29,428.471)	(38,546.407)	(310.571)	(444.128)	(8,801.905)	(11,722.266)
Own - 2	-51,945***	-19,699**	448***	354**	-46,742***	15,591
	(7,298.290)	(9,367.195)	(112.251)	(158.015)	(9,051.500)	(11,937.049)
Own - 3	3,339	-5,795*	1,662***	1,399***	-729	54,132***
	(2,285.648)	(3,022.014)	(66.480)	(92.476)	(10,126.511)	(13,315.871)
Own - 4	2,691***	-269	882***	755***	24,910*	19,405
	(317.274)	(522.517)	(43.932)	(66.006)	(12,976.968)	(17,214.775)
Own - 5	94***	668***	11**	222***	-143,540***	276,231***
	(15.961)	(31.261)	(4.628)	(17.421)	(24,920.701)	(32,920.898)
Oth - 1	129,910***	139,666***	691**	165	-90,765***	-170,848***
	(29,440.592)	(38,532.799)	(316.760)	(439.619)	(8,898.503)	(11,641.942)
Oth - 2	45,861***	12,759	-2,147***	-2,032***	31,975***	-31,215***
	(7,311.192)	(9,360.529)	(111.731)	(158.762)	(9,068.809)	(11,916.840)
Oth - 3	4,654**	10,862***	-279***	-326***	-12,250	-62,422***
	(2,307.743)	(3,002.236)	(64.037)	(93.526)	(10,030.830)	(13,408.943)
Oth - 4	-480	1,678***	-394***	-317***	-17,553	-11,169
	(359.419)	(506.517)	(49.814)	(60.419)	(12,946.797)	(17,246.836)
Oth - 5	-318***	295***	-126***	36***	183,608***	-247,901***
	(21.601)	(23.917)	(13.203)	(5.872)	(24,870.955)	(32,968.696)
Pop Diff	35***	-19*	19**	-35***	37***	-20*
	(9.172)	(10.978)	(9.218)	(11.021)	(9.050)	(10.926)
C	1,134***	1,329***	1,134***	1,329***	1,134***	1,329***
	(68.757)	(95.020)	(68.773)	(94.623)	(67.877)	(94.303)
C Sq	-102***	-92***	-102***	-92***	-102***	-92***
	(16.468)	(22.757)	(16.471)	(22.662)	(16.257)	(22.586)
C Cb	4***	2	4***	2	4***	2
	(1.081)	(1.493)	(1.081)	(1.487)	(1.067)	(1.482)
Observations	52,800	79,200	52,800	79,200	52,800	79,200
R-squared	0.400	0.169	0.266	0.176	0.285	0.182

The knots for own percent, level, and Kendall given PopDiff > 0 are located at [0.268,0.615,0.986,2.346], [2.497,5.542,8.162,11.688], and [0.096,0.280,0.416,0.665]. Knots for other correlations are located at [0.269,0.618,0.985,2.279], [2.451,5.505,8.293,11.486], and [0.268,0.615,0.986,2.346]. Knots when PopDiff is less than or equal to zero are almost identical. Not shown are seed, country, seed-country indicators, and constants. Columns are separated for larger and weakly smaller countries.

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

Table 3.29: Revenue Gap in %'s with Splines

(Pop Diff ? 0)	Percent (>)	Percent (<=)	Level (>)	Level (<=)	Kendall (>)	Kendall (<=)
Own - 1	-2.4469***	-1.7691***	-0.0945***	-0.0240***	0.2154	0.5111***
	(0.568)	(0.553)	(0.006)	(0.006)	(0.168)	(0.166)
Own - 2	-0.5893***	-0.1186	0.0014	-0.0012	-0.4752***	0.0356
	(0.140)	(0.134)	(0.002)	(0.002)	(0.173)	(0.169)
Own - 3	0.0350	-0.1169***	0.0170***	0.0179***	-0.1428	0.4241**
	(0.044)	(0.043)	(0.001)	(0.001)	(0.194)	(0.188)
Own - 4	0.0141**	-0.0137*	0.0050***	0.0063***	0.3363	-0.4098*
	(0.007)	(0.007)	(0.001)	(0.001)	(0.248)	(0.243)
Own - 5	0.0011***	0.0028***	0.0001	0.0026***	-2.9332***	4.1929***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.476)	(0.465)
Oth - 1	2.0961***	1.5059***	0.0193***	-0.0465***	-0.4868***	-1.3667***
	(0.568)	(0.553)	(0.006)	(0.006)	(0.170)	(0.165)
Oth - 2	0.4885***	-0.0355	-0.0308***	-0.0364***	0.3972**	-0.1294
	(0.141)	(0.134)	(0.002)	(0.002)	(0.173)	(0.168)
Oth - 3	0.0267	0.1596***	-0.0007	-0.0002	0.0618	-0.4942***
	(0.044)	(0.043)	(0.001)	(0.001)	(0.192)	(0.190)
Oth - 4	-0.0039	0.0218***	-0.0053***	-0.0066***	-0.2145	0.5973**
	(0.007)	(0.007)	(0.001)	(0.001)	(0.247)	(0.244)
Oth - 5	-0.0024***	-0.0012***	-0.0009***	-0.0004***	4.0176***	-3.0631***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.475)	(0.466)
Pop Diff	-0.0023***	0.0016***	-0.0023***	0.0012***	-0.0023***	0.0016***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
C	0.0212***	0.0269***	0.0212***	0.0269***	0.0212***	0.0269***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
C Sq	-0.0026***	-0.0032***	-0.0026***	-0.0032***	-0.0026***	-0.0032***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
C Cb	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***	0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	52,800	79,200	52,800	79,200	52,800	79,200
R-squared	0.328	0.302	0.325	0.301	0.359	0.333

The knots are in the same locations as in the previous regressions. Not shown are seed, country, seed-country indicators, and constants. Columns are separated for larger and weakly smaller countries.
*** p<0.01, ** p<0.05, * p<0.1

Table 3.30: Revenue Gaps with Cubics

(Pop Diff ? 0)	Percent (>)	Percent (<=)	Level (>)	Level (<=)	Kendall (>)	Kendall (<=)
Own	2,106.09***	288.89**	668.26***	-1,706.09***	40,495.07***	81,009.17***
	(78.315)	(126.522)	(13.382)	(38.849)	(10,233.503)	(13,154.163)
Own Sq	-83.08***	200.73***	-4.81***	134.82***	-101,731.53***	-231,044.19***
	(5.137)	(12.760)	(0.122)	(2.324)	(29,290.873)	(38,075.761)
Own Cb	0.91***	-6.56***	0.01***	-2.17***	55,817.09*	293,762.26***
	(0.080)	(0.316)	(0.000)	(0.039)	(30,431.123)	(39,467.478)
Oth	-1,633.61***	889.16***	-1,935.87***	357.23***	-96,910.63***	-178,698.71***
	(96.883)	(104.018)	(29.598)	(17.868)	(10,196.780)	(13,185.445)
Oth Sq	137.61***	11.75*	105.88***	-1.15***	209,060.59***	433,534.29***
	(9.539)	(6.966)	(1.795)	(0.175)	(29,111.900)	(38,231.318)
Oth Cb	-3.07***	-0.52***	-1.60***	0.00*	-106,655.72***	-410,420.31***
	(0.235)	(0.110)	(0.031)	(0.000)	(30,307.192)	(39,574.669)
Pop Diff	33.82***	-20.14*	18.05*	-6.56	37.14***	-19.99*
	(10.097)	(11.275)	(9.864)	(11.296)	(9.338)	(11.017)
C	1,134.43***	1,329.48***	1,134.43***	1,329.48***	1,134.43***	1,329.48***
	(75.717)	(97.625)	(73.931)	(97.742)	(70.042)	(95.093)
C Sq	-102.39***	-92.23***	-102.39***	-92.23***	-102.39***	-92.23***
	(18.134)	(23.381)	(17.707)	(23.410)	(16.775)	(22.775)
C Cb	3.58***	1.73	3.58***	1.73	3.58***	1.73
	(1.190)	(1.534)	(1.162)	(1.536)	(1.101)	(1.495)
Constant	382.98*	-3,097.56***	4,243.21***	1,260.56*	6,902.04***	9,171.78***
	(226.340)	(640.151)	(228.490)	(645.532)	(220.688)	(630.195)
Observations	52,800	79,200	52,800	79,200	52,800	79,200
R-squared	0.111	0.123	0.152	0.121	0.239	0.168

Not shown are seed, country, and seed-country indicators. Columns are separated for larger and weakly smaller countries.

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

Table 3.31: Revenue Gap in %'s with Cubics

(Pop Diff ? 0)	Percent (>)	Percent (<=)	Level (>)	Level (<=)	Kendall (>)	Kendall (<=)
Own	0.017752***	-0.037069***	0.006958***	-0.041429***	-0.0242	0.7586***
	(0.002)	(0.002)	(0.000)	(0.001)	(0.197)	(0.190)
Own Sq	-0.000476***	0.005274***	-0.000051***	0.002550***	0.7046	-3.0994***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.565)	(0.550)
Own Cb	0.000003**	-0.000135***	0.000000***	-0.000038***	-1.3064**	3.8137***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.587)	(0.571)
Oth	-0.052847***	0.004722***	-0.041063***	0.002948***	-0.5090***	-1.7153***
	(0.002)	(0.002)	(0.001)	(0.000)	(0.197)	(0.191)
Oth Sq	0.004938***	0.000149	0.002140***	-0.000022***	0.3936	5.2301***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.562)	(0.553)
Oth Cb	-0.000111***	-0.000005***	-0.000031***	0.000000***	0.9287	-4.8789***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.585)	(0.572)
Pop Diff	-0.002344***	0.001562***	-0.002494***	0.001627***	-0.0023***	0.0016***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
C	0.021169***	0.026912***	0.021169***	0.026912***	0.0212***	0.0269***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
C Sq	-0.002643***	-0.003168***	-0.002643***	-0.003168***	-0.0026***	-0.0032***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
C Cb	0.000111***	0.000128***	0.000111***	0.000128***	0.0001***	0.0001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.071463***	0.026892***	0.158282***	0.128791***	0.0643***	0.0577***
	(0.004)	(0.010)	(0.004)	(0.009)	(0.004)	(0.009)
Observations	52,800	79,200	52,800	79,200	52,800	79,200
R-squared	0.068	0.062	0.167	0.143	0.218	0.193

Not shown are seed, country, and seed-country indicators. Columns are separated for larger and weakly smaller countries.

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

coefficients differ in both magnitude and occasionally in sign. We can also see the non-linearity when using a double-residual regression procedure - not shown in the paper. A kernel fit on the residuals seems to suggest that the trend is at most cubic.

The preferred specification therefore includes quadratic and cubic terms. Coefficients are given in Tables 3.30 and 3.31. Figures 3.4.1 - 3.4.6 graphically depict the estimated effects of the three correlation measures on the revenue gap in levels and percentages. A correlation value of zero is deceptive because the polynomial forces the effect to be zero when the dependent variable equals zero. However, it is actually hidden in the constant term. In any, the focus of the tables and figures is on the marginal effects of own and other deviation/correlation, i.e., the slope, rather than the level effects.

In general, the revenue gap effects under percent and level deviations are similar, as evidenced by Figures 3.4.1 through 3.4.4. As own deviation increases, moving from left to right in the graphs, revenue gaps for larger countries increase although at a decreasing rate. The effects appear fairly close to being quadratic as is indicated by the relatively small cubic coefficients on own correlation. For smaller countries, the effects of decreasing own correlation is highly cubic and is actually negative at higher correlations. Comparatively, decreasing own correlation creates much stronger pro-preferential regime effects for countries with greater relative population. Decreasing other correlation almost always has a negative impact on the revenue difference for larger countries. At high levels of correlation, small decreases in correlation decrease the revenue difference. At lower levels of correlations, this flips. For smaller countries, decreasing correlation unambiguously increases the revenue difference.

Figure 3.4.1: Percent Correlation and the Rev Gap

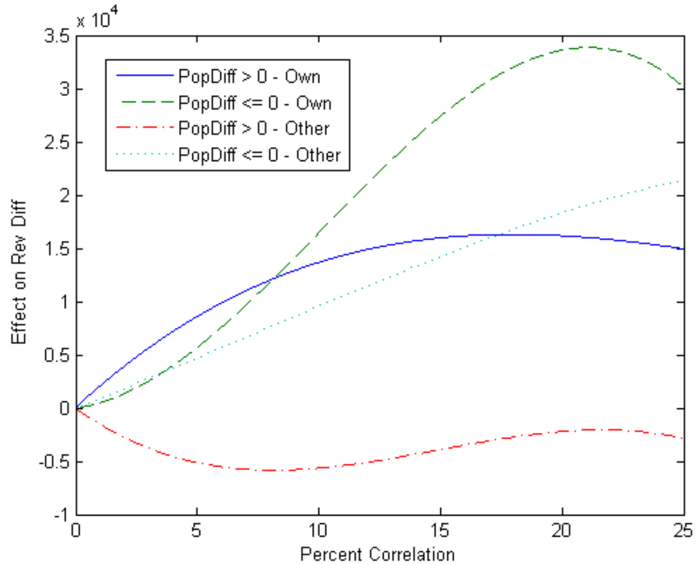


Figure 3.4.2: Percent Correlation and the Rev Gap in %

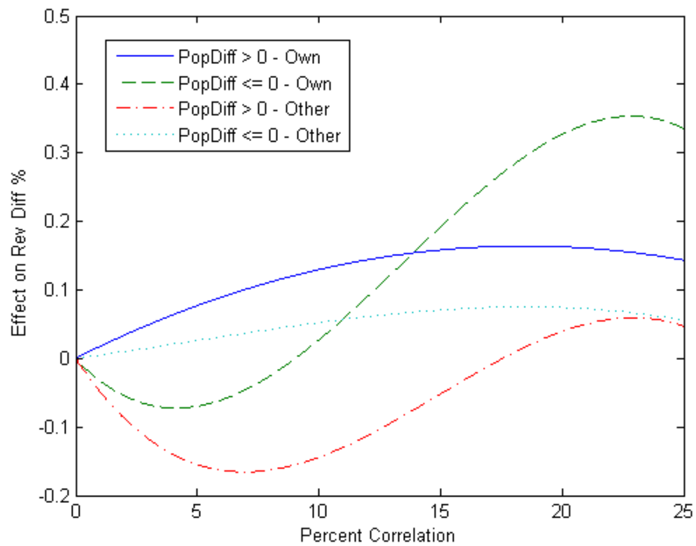


Figure 3.4.3: Level Correlation and the Rev Gap

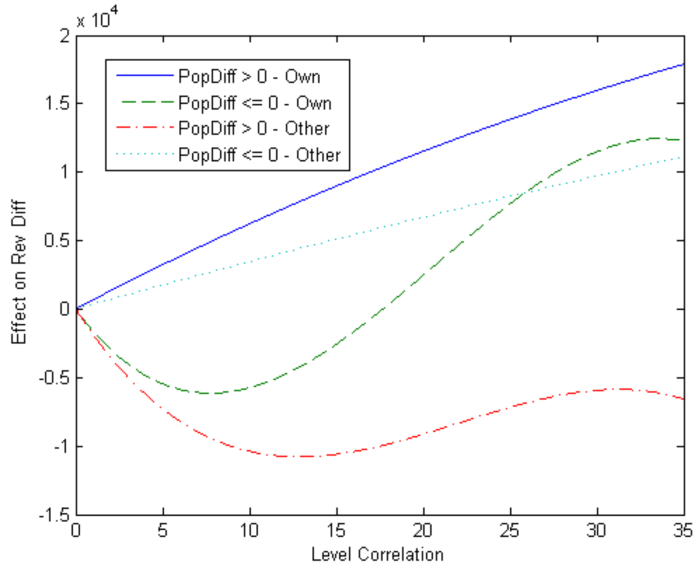


Figure 3.4.4: Level Correlation and the Rev Gap in %

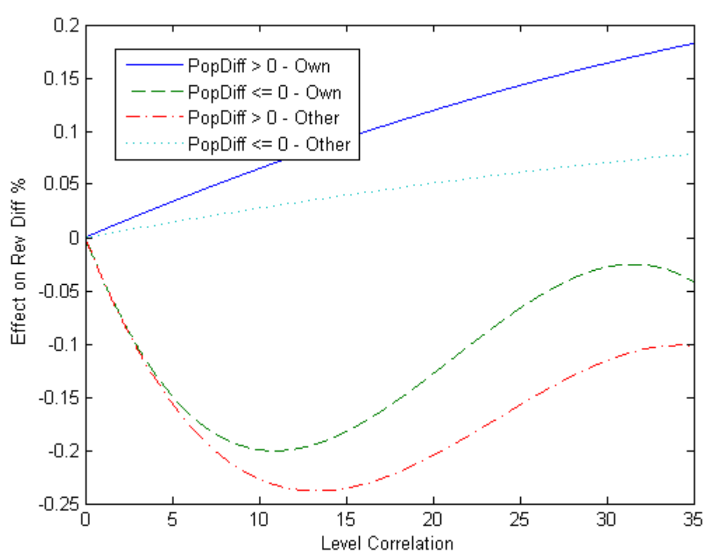


Figure 3.4.5: Kendall's Tau and the Rev Gap

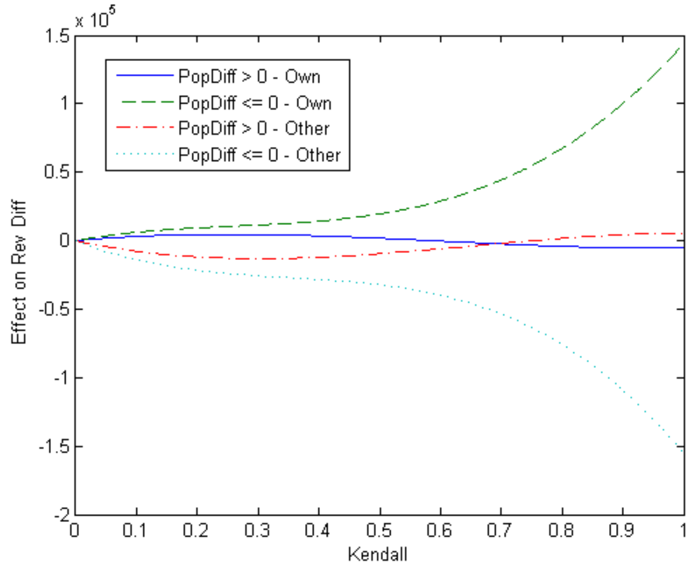
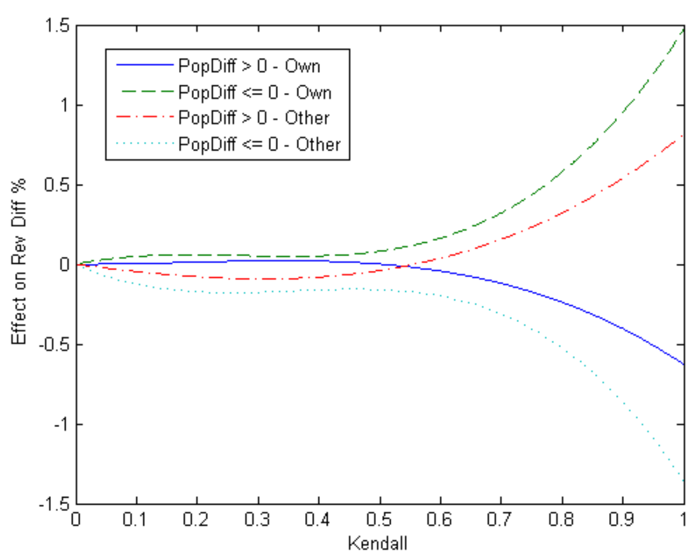


Figure 3.4.6: Kendall's Tau and the Rev Gap in %



Figures 3.4.5 and 3.4.6 graphically depict results from the Kendall's Tau correlation regressions. Recall that less correlation corresponds with moving from right to left on the graphs, i.e., Kendall's Tau decreases from 1 to 0. Decreasing own correlation tends to increase (decrease) the revenue gap for more (less) populous countries. The overall effects, however, are negative and positive, respectively. Decreasing other correlation has the opposite effect. The overall effect is positive for larger countries but decreases the revenue gap as correlation decreases. The overall effect is negative for less populous countries but increases the revenue gap as correlation decreases. The graphs are much smoother under Kendall's Tau correlation due to the direct relationship between ρ and Kendall's Tau.

If we take own and other correlations to be equal, then the combined effects of correlation are given by Figures 3.5.1 - 3.5.3. In these figures, the coefficients are summed and describe the effect of correlation if both own and other levels move together. For joint movements in own and other correlation, decreases in correlation increase revenue gaps on average and make the preferential regime more likely to be revenue-dominant.

3.5 Conclusion

This paper addressed the preferential/non-preferential question by analyzing the impact of profit correlation on the revenue-optimal regime choice. While Niu [2013b] only looked at perfect correlation and perfect non-correlation, this paper extended the model to encompass imperfect profit correlation while incorporating three aspects from the international literature. This updated, stochastic model is more realistic as it uses a Pareto distribution to represent firms as well as including randomness. Niu [2013b] found that the preferential regime was revenue-dominant in cases of asymmetry under perfectly non-correlated profits.

Under perfectly correlated profits, the preferential regime was never revenue-dominant. However, it was unclear as to the nature of the effect in between the two extreme correlations. The effect of profit correlation may be smooth or it could have been discontinuous.

Overall, results under the three measures of deviation/correlation are fairly consistent so I will focus on percent deviation. In general, changes in own versus other percent deviation have opposite trends. For a larger country, the effect of own percent deviation is initially negative at high levels of correlation but becomes positive at lower levels of correlation. This trend flips for other percent deviation. Own correlation has a greater, more positive effect on the revenue gap than other correlation. For a smaller country, these trends are the same except that other correlation is now stronger than own correlation. This makes sense as smaller countries' biggest gains from trade come via the attraction of foreign firms. If profit correlation in the larger country decreases, this results in a bigger profit differential, i.e., firms become more polarized towards either staying or moving. The smaller country can then exploit this larger wedge with a higher foreign tax.

Intuitively, less correlation decreases the negative effects of tax competition because the two tax bases become more segmented. In comparing the preferential to the non-preferential regime, the main drawback was the negative revenue effects of increased competition. All else being equal, setting specific tax rates for each base should generate greater revenue. Less competition, stemming from a decrease in the relative proportion of marginal, tax sensitive firms, therefore decreases the negative competition effect and makes preferential taxation revenue-dominant. In general, the simulations showed that decreases in profit correlation do in fact tend to increase the revenue gap. This implies that countries with lower levels of profit correlation would generate more per-country revenues if both used preferential regimes. Conversely, countries with low levels of profit correlation would generate more per-country revenues if both used non-preferential regimes.

In addition, the simulations showed that there were instances under high levels of profit correlation and zero moving costs where the preferential regime was revenue dominant. This result runs counter to the perfect mobility result from Niu [2013b] in which the revenue gap was equal to zero for $c = 0$. It is likely that this disparity is caused by the presence of randomness in these simulations. The Niu [2013b] result required that the underlying distributions be identical. While the distributions are identical in this paper, randomness was introduced due to the sampling and relatively small number of draws. For larger samples, this zero moving cost result may still apply.

The results of this paper suggest that in addition to looking at asymmetries in productivity and population, the degree of profit correlation is also an important factor in determining the revenue-optimal, bilateral regime choice. This has significant policy implications if we assume that among EU and OECD countries, country pairings are likely to have higher degrees of profit correlation than pairings between EU and developing countries. Additionally, the model also speaks towards the interplay between regime choice and ease of profit shifting. Notice that profit correlation can be interpreted as a proxy for the costliness of profit shifting between multinationals. If profit shifting was costless, then differences in the profitability of a firm across different countries would be irrelevant. Profits could simply be transferred to another jurisdiction so effectively, profits in all countries are the same. This is equivalent to assuming perfect profit correlation. As profit shifting becomes costlier, profit correlation would then decrease as location now plays a larger role in determining the level of gross profits. The results therefore suggest that preferential tax regimes are less likely to be revenue-dominant when multinational firms' abilities to profit shift are high.

Figure 3.5.1: Joint Own/Other Percent Correlation and the Rev Gap

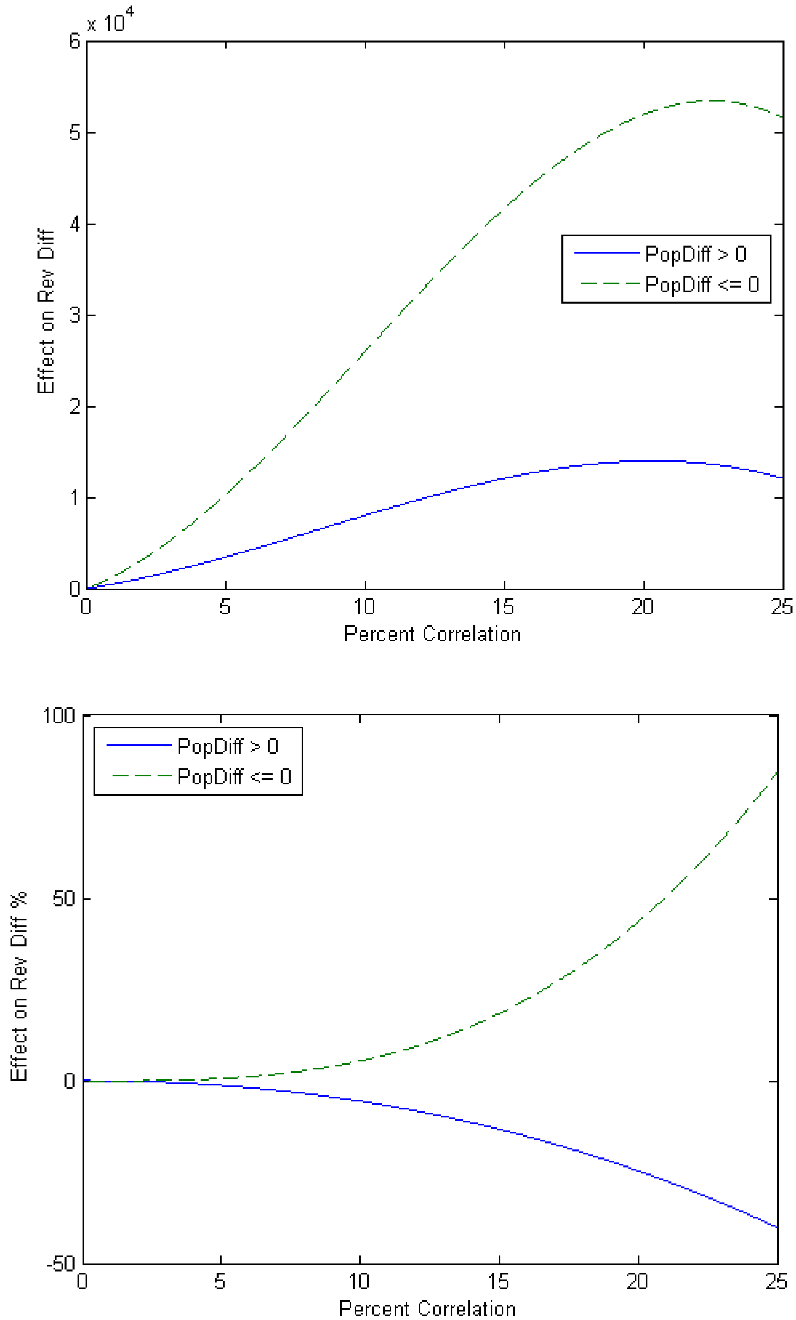


Figure 3.5.2: Joint Own/Other Level Correlation and the Rev Gap

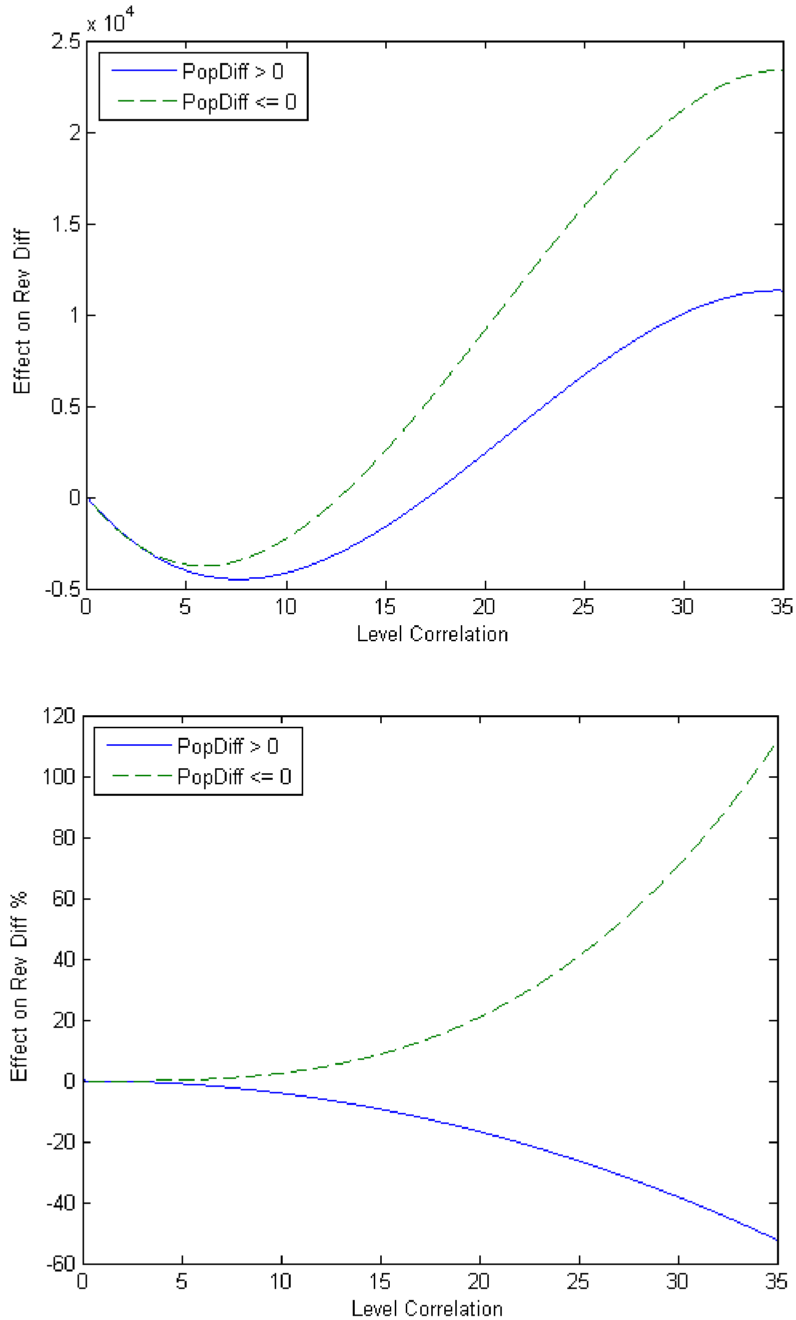
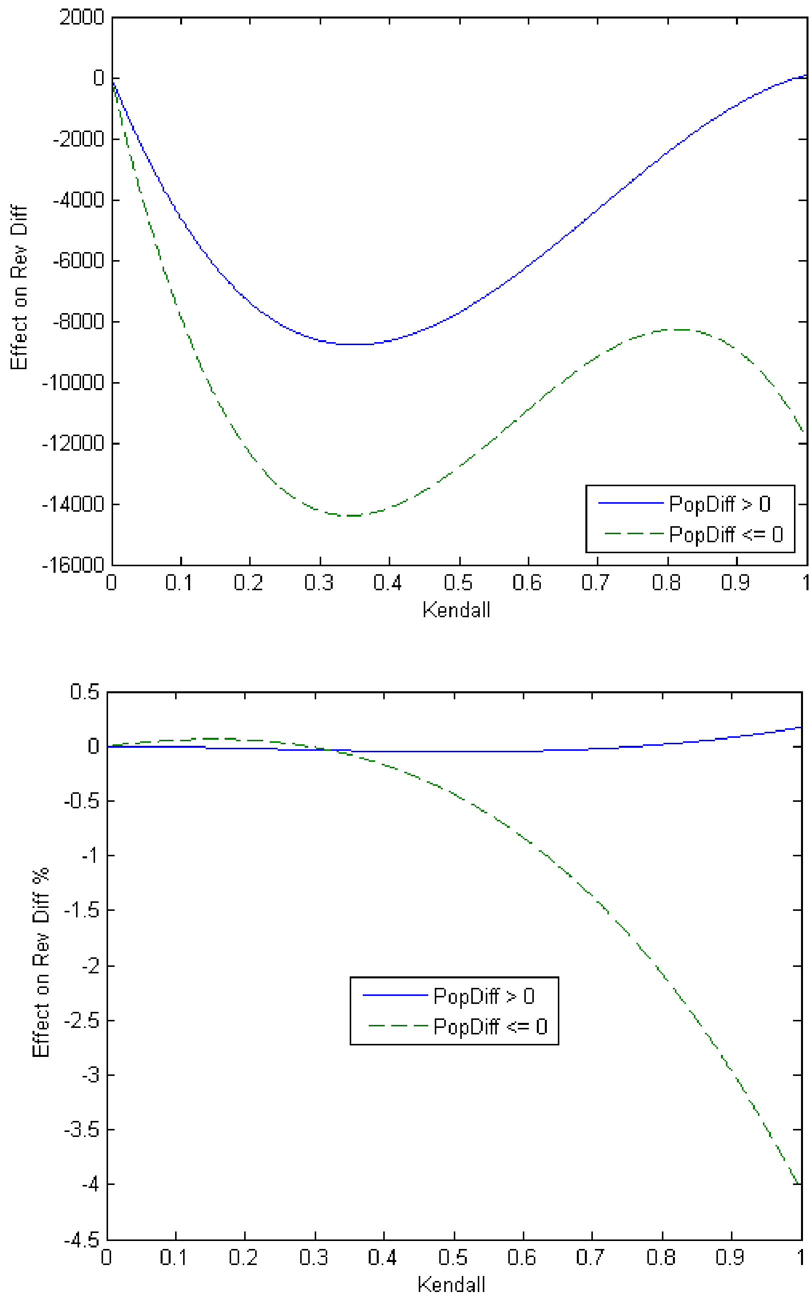


Figure 3.5.3: Joint Own/Other Kendall's Tau and the Rev Gap



APPENDICES

APPENDIX A
Simulation Equations

Households select where they want to purchase the good and whether or not to redeem by maximizing the total net benefit, U. For household i residing in city j, the total net benefit of purchasing in city k and not redeeming is given by:

$$U_{ij}^{\text{NR}} = \begin{cases} V_i - (1 + \tau_{Ind})\, p_k - (|k - j|)\, c_i^T & \text{if } k = \{1, 2\} \\ V_i - (1 + \tau_{Mich})\, p_k - 10 - (|k - j|)\, c_i^T & \text{if } k = \{3, 4\} \end{cases} \tag{A1}$$

Only purchases made in Michigan, i.e., Cities 3 and 4, are subject to the additional deposit. If the household chooses to redeem, U_{ij} increases by the ten cent deposit but decreases by the hassle cost of redemption. I refer to this as the net redemption value $10 - c_i^R$. By assumption, the travel cost associated with redemption is equal to zero for Michigan households. Michiganders will therefore choose to redeem if the net redemption value is weakly positive. The proportion of the Michigan population that redeems is given by $pr(c_i^R \leq 10) = F^R(10)$. For Indiana households, redemption results in an additional travel cost unless the household is already purchasing in Michigan. Unless the household is already purchasing in City 4, assume that it redeems in City 3.

$$U_{ij}^{\mathrm{R}} = \begin{cases} U_{ij}^{\mathrm{NR}} + \left(10 - c_i^R\right) - (3 - j)\, c_i^T & \text{if } j = \{1, 2\} \text{ and } k = \{1, 2\} \\ U_{ij}^{\mathrm{NR}} + \left(10 - c_i^R\right) & \text{if } j = \{3, 4\} \text{ and } k = \{1, 2\} \\ U_{ij}^{\mathrm{NR}} + \left(10 - c_i^R\right) & \text{if } k = \{3, 4\} \end{cases} \tag{A2}$$

In the simulations, I assumed a θ of 1.01 where θ is defined as the ratio of $c^{T,purchases}$ to $c^{T,redemptions}$. This differentiated between the travel cost incurred when a household purchases in another city versus the travel cost incurred for redemptions. More importantly, it allowed households to break ties between a number of different choices. In regards to the previous total net benefit equations, assume that c^T is now the travel cost associated with redeeming. The $(|k - j|)\, c_i^T$ term in (A1) is multiplied by θ to reflect travel costs for purchases. Equation (A2) remains unchanged.

For firms/cities, the goal is to maximize profits. Because costs are assumed to be zero this is equivalent to maximizing revenues. In the simulations, I first calculate the total net benefits of all households for each of the eight purchase-redemption combinations given prices in the four cities. Households select the option with the highest positive value; households can choose to not purchase the good. To calculate revenues in each city, I multiply the total number of households purchasing in city j by p_j. Cities optimize by setting the price that maximizes their revenue taking as given the prices in other competing cities.

APPENDIX B
Retail Price Data

Sampling wise, I first identified those cities within 20 miles of the Michigan border. This created a list of approximately 40 cities between the three states. For each of these border cities, I had a target goal of 15 retail observations between the different types of retailers with an emphasis on supermarkets, grocery stores, and liquor stores as well as retailers located within five miles of the border. However, I was only able to do so in less than ten of the border cities due to the fact the large majority these border cities did not have enough population to have such retail coverage. For cities located between 20 and 50 miles from the border, I had a target goal of five retail observations. Cities further than 50 miles had a target of only one or two observations. The list of retailers was collected using Google Maps' spatial search function with the search terms “grocery”, “beer”, and “pharmacy”.

During the actual collection of prices, there was a significant amount of attrition. Nearly half of the retailers contacted via phone either did not answer (no longer in business, busy, not present) or refused to participate (company/store policy, busy, fear of competition). In these cases, I would re-sample and try to contact another store. I also collected pricing data in-person. Three trips were taken which focused heavily on the border regions near Toledo, Ohio and South Bend, Indiana. These trips were also used to fill in attrition gaps in smaller towns and cities near the border. Attrition was also present during my in-person visits (no longer in business, missing/unknown prices, refusal to participate).

APPENDIX C
Proof of $t \geq \tau \geq T$

Recall that base b is assumed to be less responsive to tax rate changes than base B. This implies that for any $t = T$, ε_t^b is weakly greater, i.e., less negative, than ε_T^B. From (2.2.2) and (2.2.3), we know that the optimal preferential rates occur when both bases are unit elastic. Therefore, it must be true that a weakly higher rate is needed to make ε_t^b unit elastic. At the optimum rates, it must be that $t \geq T$.

Under the non-preferential regime, (2.2.4) must hold at the revenue maximizing uniform rate. Because b and B are positive, it must be that $(1 + \varepsilon_\tau^b)$ and $-(1 + \varepsilon_\tau^B)$ are of the same sign. Alternatively, both bases could be unit elastic at the same rate but this cannot be true given our assumption. Therefore, it must be that $\varepsilon_\tau^b \geq -1 \geq \varepsilon_\tau^B$ or $\varepsilon_\tau^b \leq -1 \leq \varepsilon_\tau^B$. Due to the fact that ε_τ^b is weakly greater than ε_τ^B by assumption, it must be that $\varepsilon_\tau^b \geq -1 \geq \varepsilon_\tau^B$ is true under the non-preferential regime.

Under the preferential regime, we know that at the revenue maximizing rates both elasticities are unit elastic such that $t \geq T$. Under the non-preferential regime, we know that at the revenue maximizing uniform rate $\varepsilon_\tau^b \geq -1 \geq \varepsilon_\tau^B$. Thus, the uniform rate must be less than t but greater than T which implies that $t > \tau > T$. In general, $t \geq \tau \geq T$ will be true. Both bases could be equally responsive in which case the rates are equal. However, this condition requires that the revenue functions are quasi-concave with respect to the tax rate. If the revenue function has multiple peaks, then more than one tax rate would generate the same elasticity and this condition may fail.

APPENDIX D
Optimal Rates Under the Preferential Regime

From (2.3.2) and (2.3.3), define $\phi_j^D(t_j, T_k, c, \alpha)$ and $\phi_j^F(T_j, t_k, c, \alpha)$ to be the upper bounds on the integrals over π_k. Taking the derivative of R_j with respect to t_j and T_j gives us $t_j^*(T_k, f_j, c, \alpha)$ and $T_j^*(t_k, f_k, c, \alpha)$. This is equivalent to maximizing domestic and foreign revenues individually. The equations for the optimal rates under the preferential regime are given below.

$$t_j^* = \frac{\int \int_0^{\phi_j^D} \pi_j f_j\Big(\pi_j, \pi_k\Big) \; \mathrm{d}\pi_k \mathrm{d}\pi_j}{\int_0^{\bar{\pi}_j} \frac{\pi_j^2}{1-T_k} f_j\Big(\pi_j, \phi_j^D\Big) \; \mathrm{d}\pi_j} \tag{D1}$$

$$T_j^* = \frac{\int \int_0^{\phi_j^F} (\pi_j - \alpha c) f_j\Big(\pi_j, \pi_k\Big) \; \mathrm{d}\pi_k \mathrm{d}\pi_j}{\int_0^{\bar{\pi}_j} \frac{(\pi_j - \alpha c)^2}{1-t_k} f_k\Big(\pi_j, \phi_j^F\Big) \; \mathrm{d}\pi_j} \tag{D2}$$

Under the preferential regime, the optimum tax rates are independent of country size. Intuitively, N should not affect the tax rate because the marginal increases in revenue and decreases in base are both scaled up by N. As long as the increase in firm population is proportional across all firm productivities, it should not have any effect. Alternatively, we can express the tax rates as (D3) and (D4). Notice that the N terms drop out when under the fraction.

$$t_j^* = \left(\frac{R_j^D}{N_j}\right)^{\frac{1}{2}} \left(\int\limits_0^{\bar{\pi}_j} \frac{\pi_j^2}{1 - T_k} f_j\left(\pi_j, \phi_j^D\right)\, \mathrm{d}\pi_j\right)^{-\frac{1}{2}} \tag{D3}$$

$$T_j^* = \left(\frac{R_j^F}{N_k}\right)^{\frac{1}{2}} \left(\int\limits_0^{\bar{\pi}_j} \frac{(\pi_j - \alpha c)^2}{1 - t_k} f_k\left(\pi_j, \phi_j^F\right)\, \mathrm{d}\pi_j\right)^{-\frac{1}{2}} \tag{D4}$$

APPENDIX E
Optimal Rates Under the Non-Preferential Regime

With $t_j = T_j$, the optimum uniform tax rate τ_j now affects both the domestic and foreign tax revenue.

$$\tau_j^* = \left(R_j^D + R_j^F\right)^{\frac{1}{2}} \Bigg(\int_0^{\bar{\pi}_j} \frac{\pi_j^2}{1-\tau_k} \left(N_j f_j\left(\pi_j, \phi_j^D\right)\right)$$

$$- \frac{\left(\pi_j - \alpha c\right)^2}{1-\tau_k} \left(N_k f_k\left(\pi_j, \phi_j^F\right)\right) \ \mathrm{d}\pi_j \Bigg)^{-\frac{1}{2}} \qquad \text{(E1)}$$

Under the non-preferential regime, the optimum uniform tax rate does depend on the size of the countries. As opposed to the case of the preferential regime where a single country's domestic and foreign bases were independent of each other, the uniform tax rate now affects both bases simultaneously. An increase in τ_j causes a marginal increase in revenue from existing domestic and foreign firms, and a marginal decrease in revenue from those leaving the two bases. The domestic base is scaled by the N_j term while the foreign base is scaled by the N_k term. Notice, however, that if both N_j and N_k change proportionally, τ_j is unaffected.

APPENDIX F
Ch. II Numerical Details

I use MATLAB to analyze my numerical model. The basic outline is listed below.

1. Calculate Country A's revenue function taking as given the tax rates of Country B.
2. Find the argmax(es) of the revenue function.
3. Repeat steps 1 and 2 for each possible combination of Country B's tax rates (t_B,T_B) to complete Country A's best response function.
4. Repeat steps 1 - 3 for Country B's best response function.
5. Identify pure strategy equilibrium/equilibria by searching for corresponding entries in the two best responses.

From 2.3.1, I can explicitly write down the equations for the population of firms that are moving or not moving as an integral over the distribution of firms' profits. Multiplying the densities by the profit times the tax rates gives me the domestic/foreign revenues for a given country. Because firms are distributed uniformly the density is straightforward. However, the general equation is highly piece-wise so a number of conditional statements are needed to correctly calculate the revenue under different sets of parameters. A's revenue function (matrix), taking as given B's current tax rates, is calculated by running through A's revenue equations for each possible tax rate. Tax rate(s) that maximize revenue given B's current tax rates are recorded as best responses. The loop then repeats but for the next combination of B's tax rates. If the countries are symmetric, then A's and B's best response functions (matrices) are identical. If the countries are asymmetric, then I repeat this process under the new set of parameters where A and B are reversed. Pure strategy equilibria are calculated by looking for intersection(s) in the best response functions. It is possible to use other, numerically faster methods to identify the equilibria. However, corner solutions as well as multiple equilibria are of particular concern in this literature and type of model. As such, I felt it more appropriate to identify the entire range of equilibria using a more rudimentary but thorough method.

To calculate the TNP under each equilibrium, the tax rates are plugged back into the previous revenue equations. Firms move according to the equilibrium tax rates but the amount of revenue collected from each firm is set to one, i.e., tax collection is set at 100%, to capture the profits of all firms. The moving cost is deducted from the profits of those foreign firms. The TNP loss is calculated by comparing the TNP under each equilibrium against the TNP when tax rates are set to zero. Given the computing time required, the majority of the programs are run on the University of Michigan CAC NYX/FLUX computing cluster.

For the main simulations, $\bar{\pi}_j = \{1, 1.25, 1.5, 1.75, 2\}$ and likewise for the population multipliers N_j. For moving costs and moving cost deductibility, I consider the values $c = \{0, 0.1, 0.2, 0.3\}$ and $\alpha = \{0, 0.25, 0.5, 0.75, 1\}$. Every combination of productivity, population, moving cost, and cost deductibility is then considered in this range of parameter values. For the more specific productivity and population asymmetry graphs shown in Figures 2.4.5 - 2.4.7, I simulate cases for $\bar{\pi}$ and N between 1 and 100 in smaller increments to capture specifically targeted effects.

APPENDIX G
Ch. III Numerical Details

To generate the samples from the bivariate HRT, I first created two vectors of random numbers drawn uniformly between 0 and 1. The first vector houses the CDF values on the π_{home} marginal distribution. Inverting the Pareto marginal allowed me to back out the corresponding π_{home} value that would have generated the corresponding CDF value. The second vector houses the conditional CDF values, i.e., $Pr(\pi_{\text{abroad}}|\pi_{\text{home}})$. Because the HRT copula and marginals are invertible, this allowed me to back out the second profit term conditional on the first, for different values of ρ and other parameters. The same process was then repeated from the perspective of the other country.

The grid-search begins by arbitrarily choosing a starting tax rate for the other country - assume that this is Country B. Country A proceeds to narrow down the best response by going through three meshes of increasing fineness: 0.05 (5%), 0.001 (0.1%), and 0.0001 (0.01%). At each mesh, Country A finds the tax rate that maximizes revenue. It then narrows the scope of possible rates to be $\pm$ the fineness step and repeats the search at the next level of fineness. After Country A nails down the optimal response, Country B repeats this process taking as given A's response. This continues until an equilibrium is reached. In cases where a pure strategy Nash Equilibrium was not found, an equilibrium loop was identified, i.e., a closed cycle of reaction-counter-reaction. The magnitude of different rates within these loops were typically $\pm 0.01\%$. As such, the average of the rates within these loops were used to approximate a pure strategy equilibrium. In these simulations, this is caused by the relatively small number of draws per country. Higher numbers of firms would alleviate this issue but at the cost of much greater computing needs.

Bibliography

Michigan bottle deposit law frequently asked questions. Technical report, Dept. of Treasury, April 2010. URL `http://www.michigan.gov/deq`.

American Beverage Assocation v. Rick Snyder, Bill Schuette, and Andrew Dillon, 2011.

State beer excise tax rates, 2009-2013. Technical report, Tax Foundation, 2013. URL `http://taxfoundation.org/`.

David Agrawal. Games within borders: Are geographically differentiated taxes optimal? *Intern*, 4:574–597, 2012.

Bevin Ashenmiller. Externalities from recycling laws: Evidence from crime rates. *American Law and Economics Review*, 12:245–261, 2010.

Bevin Ashenmiller. The effect of bottle laws on income: New empirical results. *The American Economic Review*, 101, 2011.

Robert Axtell. Zipf distribution of u.s. firm sizes. *Science*, 293:1818–1820, 2001.

Richard Baldwin and Paul Krugman. Agglomeration, integration and tax harmonisation. *European Economic Review*, 48:1–23, 2004.

Richard Baldwin and Toshihiro Okubo. Tax competition with heterogeneous firms. February 2009 Working Paper.

Leon Bettendorf and Hendrik Vrijburg. Asymmetric corporate tax competition with two tax instruments. In *IIPF Conference Uppsala, Sweden*, 2010.

Jagdish Bhagwati and Bent Hansen. A theoretical analysis of smuggling. *The Quarterly Journal of Economics*, 87 (2):172–187, 1973.

Tilman Börgers. Iterated elimination of dominates strategies in a bertrand-edgeworth model. *Review of Economic Studies*, 59:163–176, 1992.

Sam Bucovetsky and Andreas Haufler. Preferential tax regimes with asymmetric countries. *National Tax Journal*, 60(4):789–795, December 2007.

John Burbidge, Katherine Cuff, and John Leach. Tax competition with heterogeneous firms. *Journal of Public Economics*, 90:533–549, 2006.

Colin Cameron, Jonah Gelbach, and Douglas Miller. Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90:414–427, 2008.

Ronald Davies and Carsten Eckel. Tax competition for heterogeneous firms with endogenous entry. *American Economics Journal*, 2010.

Michael Devereux, Ben Lockwood, and Michela Redoano. Do countries compete over corporate tax rates? *Journal of Public Economics*, 92:5–6, 2008.

Julian di Giovanni, Andrei Levchenko, and Romain Ranciere. Power laws in firm size and openness to trade. *Journal of International Economics*, 2011.

Ian M. Dobbs. Litter and waste management: Disposal versus user charges. *The Canadian Journal of Economics*, 24:221–227, 1991.

Wolfgang Eggert and Alfons J. Weichenrieder. On the economics of bottle deposits. *CESifo Working Paper Series No. 1243*, 2004.

Charles Engel and John H. Rogers. How wide is the border? *The American Economic Review*, 86:1112–1125, 1996.

Yoshi Fujiwara, Corrado Di Guilmi, Hideaki Aoyama, Mauro Gallegati, and Wataru Souma. Do pareto-zipf and gibrat laws hold true? an analysis with european firms. *Physica A*, 335:197–216, 2004.

Don Fullerton and Ann Wolverton. Two generalizations of a deposit-refund system. *The American Economic Review*, 90(2):238–242, 2000.

Don Fullerton and Wenbo Wu. Policies for green design. *Journal of Environmental Economics and Management*, 36:131–148, 1998.

Philipp Genschel and Peter Schwarz. Tax competition: a literature review. *Socio-Economics Review*, 9:339–370, 2011.

Alexander Haupt and Wolfgang Peters. Restricting preferential tax regimes to avoid harmful tax competition. *Regional Science and Urban Economics*, 35:493–507, 2005.

Eckhard Janeba and Wolfgang Peters. Tax evasion, tax competition and the gains from nondiscrimination: The case of interest taxation in europe. *The Economic Journal*, 109: 93–101, 1999.

Eckhard Janeba and Michael Smart. Is targeted tax competition less harmful than its remedies? *International Tax and Public Finace*, 10:259–280, 2003.

Ravi Kanbur and Michael Keen. Jeux sans frontieres: Tax competition and tax coordination when countries differ in size. *The American Economic Review*, 83 (4):877–892, 1993.

Michael Keen. Preferential regimes can make tax competition less harmful. *National Tax Journal*, 54:757–762, 2001.

K.K. Lewis. Trying to explain home bias in equities and consumption. *Journal of Economic Literature*, 37:571–608, 1999.

Michael Lovenheim. How far to the border?: The extent and impact of cross-border casual cigarette smuggling. *National Tax Journal*, 61:7–33, 2008.

K.V. Mardia. Multivariate pareto distributions. *Annals of Mathematical Statistics*, 33:1008–1015, 1962.

David Merriman. The micro-geography of tax avoidance: Evidence from littered cigarette packs in chiago. *American Economic Journal: Economic Policy*, 2:61–84, 2010.

Steeve Mongrain, John D. Wilson, and Nicolas Marceau. Why do most countries set high tax rates on capital? *Journal of Int*, 80:249–259, 2010.

Ben Niu. Preferential corporate taxation and profit correlation under a bivariate pareto distribution. 2013a.

Ben Niu. Equilibria and location choice in corporate tax regimes. 2013b.

Wallace E. Oates. *Fiscal Federalism.* Harcourt Brace Jovanovich, NY, 1972.

Kosuke Oshima. Preferential tax regimes and asymmetries of countries. *Journal of the UMDS - Information, Economics & Management Science*, 18:1–11, 2009.

Kosuke Oshima. Preferential tax regimes and agglomeration economies. *Journal of the UMDS - Information, Economics & Management Science*, 19:49–59, 2010.

P. Osmundsen, KP Hagen, and G. Schjelderup. Internationally mobile firms and tax policy. *Journal of International Economics*, 45:97–113, 1998.

Peter Spendelow. Fact sheet: The oregon bottle bill. Technical report, State of Oregon Dept. of Environmental Quality, October 2004. URL `www.deq.state.or.us`.

Charles Tiebout. A pure theory of local expenditures. *Journal of Political Economy*, 64: 416–424, 1956.

US-CensusBureau. Statistics of us businesses. Technical report, 2009.

W. Kip Viscusi, Joel Huber, and Jason Bell. Discontinuous behavioral responses to recycling laws and plastic water bottle deposits. *American Law and Economics Review*, 15:110–155, 2013.

Henrik Vrijburg and Ruud A. de Mooij. Enhanced coorporation in an asymmetric model of tax competition. 2010.

John D. Wilson. A theory of interregional tax competition. *Journal of Urban Economics*, 19:296–315, 1986.

John D. Wilson. Theories of tax competition. *National Tax Journal*, 52:269–304, 1999.

John D. Wilson and Steeve Mongrain. Tax competition with and without tax discrimination against domestic firms. June forthcoming.

George Zodrow and Peter Mieszkowski. Pigou, tiebout, property taxation, and the under-provision of local public goods. *Journal of Urban Economics*, 19:356–370, 1986.