Wholesale Market Power and the Limited Profitability of Retail Zone Pricing

Eugenio J. Miravete* Katja Seim† Jeff Thurk‡

January 2019

PRELIMINARY AND INCOMPLETE – DO NOT CITE

Abstract

We present a general analysis of third-degree price discrimination in vertical markets. We show that downstream third-degree price discrimination impacts upstream demand, thereby leading the upstream firm to modify its pricing decision which in turn impacts downstream costs. We show this response not only reduces the profitability of downstream third-degree price discrimination, it also means the relative demand curvature condition of Aguirre, Cowan, & Vickers (2010) is no longer sufficient for price discrimination to increase welfare. Numerical experiments indicate the strategic behavior of upstream firms can have a large adverse effect on downstream profits and consumer welfare.

Keywords: Third-degree price discrimination, vertical markets, uniform vs. zone pricing, returns to pricing complexity, pass-through.

JEL Codes: D42, D63, L43, L66

* The University of Texas at Austin, Department of Economics, 2225 Speedway Stop 3100, Austin, Texas 78712-0301; Centre for Competition Policy/UEA; and CEPR, London, UK. Phone: 512-232-1718. Fax: 512-471-3510. E-mail: miravete@eco.utexas.edu; http://www.eugeniomiravete.com

† Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6372; and NBER. E-mail: kseim@wharton.upenn.edu; https://bepp.wharton.upenn.edu/profile/712/

‡ University of Notre Dame, Department of Economics, Notre Dame, IN 46556. E-mail: jthurk@nd.edu; http://www.nd.edu/jthurk/
1 Introduction

Zone pricing is a third degree price discrimination strategy commonly used in retailing where products are sold at different prices in separate local markets. Third degree price discrimination (3DPD) increases retailers’ profits if preferences are sufficiently heterogeneous across locations. Robinson (1933, §15.5) initiated a several decades-long research agenda aimed to show that 3DPD also enhances welfare if it increases aggregate sales over the sales level under uniform pricing. She hints at the relative curvature of demands across local markets as the key driver of this overall sale increase.\footnote{Schmalensee (1981) proved that welfare increases if total output expands under 3DPD for a constant marginal cost monopolist with independent local demands across markets. Varian (1985) generalized this result to the case of interdependent demands and nondecreasing marginal costs. Finally, Schwartz (1990) further generalized the same result for any cost function that depended on total output alone.} Aguirre, Cowan and Vickers (2010) is the most recent and general analysis of this long-standing question. They show that total output expands and welfare increases with price discrimination when the inverse demand in the low-price market is more convex and the price difference between uniform and price discrimination is small.

A key assumption of all these theoretical papers on 3DPD is that retail unit costs remain constant, thus implicitly assuming a competitive wholesale market. In the present work we challenge this assumption and allow for upstream firms with market power able to set their optimal wholesale price in response to a downstream retailer’s decision to price discriminate. Our analysis shows that the upstream firm generally increases the wholesale price in response to the retailer’s price discrimination decision, that this increase in retailer costs reduces the profitability of 3DPD, resulting in a smaller optimal price difference across local markets. More importantly though, our analysis shows that the relative convexity of demand condition is not longer sufficient for 3DPD to increase welfare if wholesalers have market power as the retail cost increase may lead to a reduction of aggregate sales.

All these results indicate that the incentives to price discriminate across locations are much diminished if wholesale firms enjoy some market power. This has important consequences for the empirical analysis of 3DPD. It is a well established fact that profits increase with price discrimination relative to uniform pricing. However, the empirical evidence available documents that firms frequently favor nearly-uniform pricing across local markets. Indeed, many recent studies claim that retail chains engage less in zone pricing than local demographic differences suggest they should.\footnote{Della Vigna and Gentzkow (2017) use the Nielsen-Kilts retail panel to evaluate sales of 10 products across 20,000 stores by 73 chains. With such vast amount of information they conclude that retail chains give up 7%. Two additional works confirm the use of nearly uniform pricing: Adams and Williams (2018) in the context of duopoly hardware chains and Hitcsh, Hortacsu and Lin (2017) using over 50,000 products sold in retail chains.} Is this practice of near-uniform pricing somewhat reprehensible? Are retail chains pricing decisions damaging to consumers? But most importantly, why don’t retail firms want to capture the rents of market segmentation? To reconcile theory and business practice and justify why firms do not price discriminate more often, this evidence is currently interpreted as supporting the existence of managerial decision-making costs, an argument put forward by Bloom and Van Reenen (2007).
The estimates of foregone profits of not engaging in 3DPD are therefore interpreted as equilibrium estimates of these managerial decision-making costs.

There are, of course, alternatives explanations that, without necessarily invalidating the managerial cost hypothesis, might qualify its importance. Many of them are relevant for the retail chain industry. First, the ability of firms to price discriminate is limited by the existence of arbitrage costs that are routinely assumed to be low enough although there is no evidence supporting it. Second, Corts (1998), Holmes (1989), and Thisse and Vives (1988) show that strategic price discrimination considerations may lead to lower profits with 3DPD in the oligopolistic environments where retail chains operate. This hypothesis was confirmed empirically by Dobson and Waterson (2005). And third, our research focus, wholesalers are not necessarily price takers as unanimously assumed. While 3DPD appears profitable if retail costs remain unchanged, its profitability might be substantially diminished if a supplier with market power changes wholesale prices in response to retailers’ decisions to price discriminate. The reason is that the aggregate wholesale demand changes with retail 3DPD relative to uniform pricing and therefore a supplier with market power adjusts wholesale price to remain on its elastic demand region.

The predictions of our model are of particular relevance for the empirical analysis. Following theory closely, empirical models routinely consider wholesale prices as invariant to retailers’ price discrimination strategies. Unfortunately, this assumption is likely violated in practice as retail chain firms sell products distributed by wholesalers with substantial market power. Retail profits of 3DPD will be lower if wholesale prices increase when retailers choose to price discriminate. Thus, ignoring the market power of wholesalers will overestimate the managerial decision-making costs. Furthermore, if the price spread from price discrimination is smaller than anticipated in a theoretical framework with constant wholesale prices, the mounting evidence of chain stores engaging in nearly-uniform pricing might be simply confirming that wholesalers have substantial market power and that they could capture a large share of the incremental downstream profits from retail price discrimination.

We address the ability of a wholesale monopolists to revise its price in response to the price discrimination strategy of retail monopolist with three differentiated but complementary approaches: a theoretical framework, the numerical solution of a stylized model, and a counterfactual retail pricing analysis of the Pennsylvania liquor industry. Each serves a specific purpose: the theoretical model identifies general conditions for a wholesale price increase in response to retail price discrimination and shows that aggregate sales and welfare may not increase even if the usual relative curvature conditions of local demands hold; the numerical analysis shows the magnitude of the retail profit reduction as a function of local demand heterogeneity (demand curvature); and the counterfactual analysis illustrates how these results also occur in multiproduct environments with upstream oligopoly wholesalers.

---

3 Dearing (2016) allows wholesale price to vary with the retailer’s price discrimination strategy but only considers two specific examples—iso-elastic and linear demands—that are not generalizable as overall output, welfare, and the relative size of wholesale and retail profits depends critically on the curvature of local demands and less so on the elasticity of demand.
First, our theoretical analysis relies on the approach of Aguirre et al. (2010) to evaluate how aggregate demand changes as a result of both, the retailer’s decision to price discriminate and the wholesaler’s decision to choose its profit maximizing price. We show that the wholesale price increases in response to the retailer’s decision to price discriminate under very general demand conditions. All that is needed is that the upstream demand elasticity is non-increasing with the retail price differential across local markets. This upward revision of the wholesale price leads to a necessary reduction in retail profits which are not accounted for when wholesale price is wrongly assumed constant. We also show that since retail profitability of 3DPD is diminished when the wholesaler revises its pricing upwards, the optimal retail price difference across local markets is also smaller than expected in a world with constant wholesale price. The reason behind this result is the higher pass-through rate in the weak than in the strong market (with weak and strong market being defined by their price level under 3DPD relative to uniform pricing) if the usual demand curvature conditions leading to an increase of sales and welfare with price discrimination hold. Lastly, the combination of these two effects, a smaller optimal price dispersion and a large enough reduction of wholesale price elasticity, may indeed reduce aggregate sales with retail price discrimination if wholesale price increases sufficiently. The usual relative curvature of local demands condition, a result that took nearly eighty years to be established, is thus no longer sufficient to ensure that 3DPD will enhance welfare.

Second, we solve numerically a model of constant-curvature local demand – Bulow and Pfleiderer (1983) – to compare predictions on aggregate sales, pricing, and profits of 3DPD relative to uniform pricing both, when wholesale price is wrongly assumed to be independent of retailer’s pricing strategy, and when the wholesaler revises its pricing decision in anticipation of the retailer’s strategy. The relative curvature of local markets is not only responsible for overall output and welfare to increase or decrease with 3DPD but also the key driver behind the reduction in retail profitability of 3DPD and the price difference across markets. This numerical analysis shows how results depend on the curvature of demand and curve heterogeneity across local markets. In particular, it shows that the upward revision of wholesale prices, the reduction of the price dispersion across local markets, and the reduction of aggregate sales is more likely the higher the pass-through ratio of demand is as well as the demand heterogeneity across local markets.

Third, our analysis of the Pennsylvania market for spirits documents how much aggregate sales and wholesale prices increase if 3DPD was allowed to segment consumers in smaller local markets; how much smaller price differences across local markets would be if wholesalers have market power and revise their pricing; and, most importantly, how much wholesalers’ pricing revisions limits the profitability of retail 3DPD. Our results suggest that upstream distillers with market power successfully reduce the profitability of retail 3DPD by 25% after revising their pricing decisions. Therefore, ignoring wholesale price responses to retail price discrimination can not only overstate the magnitude of managerial costs but also helps explaining the use of nearly uniform pricing in retailing.
To conduct these counterfactuals, we make use of the data and estimates of demand for spirits of Miravete, Seim and Thurk (2018a). We evaluate the incremental profits of zone price discrimination over uniform pricing by the Pennsylvania Liquor Control Board (PLCB). Unlike the stylized theoretical model we deal with multiple products and more than two local markets to account for 3DPD: our data includes all sales of 312 spirits in all 454 stores of a state-run monopolist, with our demand estimation accounting for cross-price substitution among all these spirits. The institutional environment of the Pennsylvania alcohol has several advantages. First, the PLCB is a retail monopolist and we can rule out strategic considerations in choosing 3DPD over uniform pricing. Second, our data covers the whole market rather than a sample of firms. Since all stores sell the same products at the same price we can ignore consumers’ store choices (Thomassen, Smith, Seiler and Schiraldi, 2017). Finally, the distilling industry is oligopolistic and wholesalers will re-optimize their prices to accommodate the PLCB’s chosen 3DPD strategy in our counterfactual analysis (Villas-Boas, 2007).

In addition to documenting that wholesalers’ pricing decisions might be ultimately responsible for retailers’ preference for uniform pricing, our counterfactual analysis also evaluates its impact on welfare in a context involving hundreds of local stores and hundreds of related products (Chintagunta, Dubé and Singh, 2003). We further account for location distribution of consumers with heterogeneous preferences and evaluate who benefits and who suffers from uniform pricing.4

The paper is organized as follows. Section 2 shows that wholesaler’s reaction to retail’s zone pricing generally results in wholesale price increase that in return reduces retailer’s incentive to price local markets differently. Section 3 make use of constant-curvature local demand functions of Bulow and Pfleiderer (1983) to solve for 3DPD with restricted constant and optimal wholesale pricing. The Appendix further elaborates on the solution of the model, its properties, and how results change with curvature of demand, and how similar demands are across markets. In section 4 we use data and estimates from a companion paper (Miravete, Seim and Thurk, 2018b) to evaluate the robustness of the predictions from the theoretical model using a more sophisticated model of consumer demand and oligopoly pricing common to the industrial organization literature and often used to inform government policy. Section 5 concludes.

4Eizenberg, Lach and Yiftach (2017) evaluate this redistribution effect in the opposite case when grocery retailers appear to engage in 3DPD to segment markets according to residents’ income levels.
2 A Simple Theoretical Framework

Consider a wholesale monopolist with a constant marginal cost $\zeta$ selling at a constant price $c$ to a monopoly retailer that distributes the good to final consumers without incurring in additional costs. The retailer may sell at a uniform price across all stores or split them into two local markets. Hence, the retailer purchases each unit at a cost $c_i$ and charges $p_i$ to consumers, with $i = w$ denoting the weak market, $i = s$ the strong market, and $i = u$ the uniform retail pricing solution where prices are the same across both local markets. Local price with discrimination relative to the uniform pricing solution characterizes a market as strong or weak, with $p^w < p^u < p^s$.

The timing of the game is as follows. The wholesale monopolist chooses upstream price $c$. A single downstream retailer observes $c$ and then decides whether or not to engage in 3DPD. If it chooses 3DPD, it may incur a fixed fee $F \geq 0$ and chooses market-level prices $(p^w, p^s)$. If not, it chooses a uniform price $p^u$ for consumers in both markets. For the numerical analysis of this section we consider the case where $F = 0$, e.g., nil managerial decision making costs, so the downstream firm always profits from 3DPD. This also focuses the analysis to the case where the upstream firm’s choice of $c$ impacts the degree to which the downstream firm price discriminates.

We build upon Aguirre et al. (2010) and adopt their partial equilibrium analysis that focuses on small price deviations from uniform pricing. We however allow the wholesale price $c$ to be endogenous. This is the key difference between our analysis and all previous theory models addressing the welfare effect of 3DPD going back to Robinson (1933). As the retailer’s decision to price discriminate determines overall sales, the wholesaler sets $c$ to maximize its profits in anticipation of the optimal prices $\{p^w, p^s\}$ under retail 3DPD, or $p^u$ with uniform pricing, respectively. Only exceptionally the wholesale price $c$ is independent of the retailer’s price discrimination choice: either the upstream industry is competitive or local demands are linear. Therefore, incentives for retail 3DPD might be smaller than normally anticipated because as the wholesaler revises its price, it might capture a larger share of the integrated industry profits.

Demand in market $i = \{w, s\}$ is $q^i(p^i)$ under 3DPD and $q^i(p^u)$ with uniform pricing. Let $\eta^i(p^i) \equiv -p^iq^u_p(p^i)/q^i(p^i)$ denote the price elasticity and $L^i(p^i, c) \equiv (p^i - c)/p^i$ be the Lerner index in each market. Next, define $\alpha(p) \equiv -p \cdot q_{pp}(p)/q_p(p)$ as the curvature of the direct demand. Similarly, the curvature of the inverse demand, $\sigma(q) \equiv -q \cdot p_{qq}(q)/p_q(q) = q(p) \cdot q_{pp}(p)/[q_p(p)]^2$, which is obtained after applying the inverse function theorem. Furthermore, notice that $\eta = \alpha/\sigma$.

The retailer’s profit maximizing price in market $i$ when the wholesaler charges $c$ is:

$$p^i(c) \in \arg\max_{p^i} \pi^i(p^i, c) = (p^i - c)q^i(p^i),$$

which requires the following condition to hold for each market:

$$\pi^i_p(p^i, c) = q^i(p) + (p^i - c)q^i_{qq}(p^i) = q^i(p)[1 - L^i(p^i, c)\eta^i(p^i)] = 0.$$


Local demand curvatures determine how much of a wholesale price increase is passed to consumers in each local market. We will show that this different pass-through rate in each local market is responsible for the reduction in the spread of local retail prices when the wholesale price increases. Totally differentiating the first order condition (2) with respect to \( p \) and \( c \) and substituting \( p^i - c = -q^i / q^i_p \) we get:

\[
\frac{dp^i}{dc} = \frac{q^i_p}{\pi^i_p} = \frac{q^i_p}{2q^i_p - q^i \cdot q^i_{pp}/q^i_p} = \frac{1}{2 - \sigma^i} > 0.
\]

Notice that local retail pass-through ratios are incomplete for \( \sigma^i < 1 \). Furthermore, the profit function is assumed to be concave in price, which is ensured if \( \sigma^i < 2 \):

\[
\pi^i_{pp}(p^i, c) = 2q^i_p(p) + (p^i - c)q^i_{pp}(p^i) = [2 - L^i \alpha^i]q^i_p(p) = [2 - \sigma^i]q^i_p(p) < 0.
\]

Next, let \( \delta \leq p^s - p^w \) denote the degree of discrimination. Without loss of generality, we make \( p^w = p \) and \( p^s = p + \delta \). Thus, total sales across local markets at \( p \) are:

\[
Q(c, \delta) = q^w[p^w(c, \delta)] + q^s[p^s(c, \delta)].
\]

The profit function of a retail monopolist that price discriminates is:

\[
\pi(p, \delta, c) = \pi^w(p, c) + \pi^s(p + \delta, c) = (p - c)Q(c, \delta) + \delta q^s[p^s(c, \delta)],
\]

which is decreasing in wholesale price both overall and for each local market, i.e., \( \pi^w(p, \delta, c) < 0 \), \( \pi^w(p, c) < 0 \), and \( \pi^s(p + \delta, c) < 0 \). For a given degree of discrimination \( \delta \), 3DPD requires solving:

\[
\pi^w_p(p, c) + \pi^s_p(p + \delta, c) = 0,
\]

with \( \delta = 0 \) characterizing the uniform pricing solution. Totally differentiating (7) leads to:

\[
(\pi^w_{pp} + \pi^s_{pp})dp + \pi^w_{pp}d\delta + (\pi^w_{pc} + \pi^s_{pc})dc = 0.
\]

\[\text{Aguirre et al. (2010)}, \text{like all previous analyses of 3DPD, assume } dc = 0 \text{ and show that prices will decrease in the weak market and increase in the strong market with discrimination relative to uniform pricing:}\]

\[
\frac{dp^w}{d\delta} = \frac{-\pi^s_{pp}}{\pi^w_{pp} + \pi^s_{pp}} < 0 \quad \text{and} \quad \frac{dp^s}{d\delta} = \frac{d(p^w + \delta)}{d\delta} = \frac{-\pi^w_{pp}}{\pi^w_{pp} + \pi^s_{pp}} > 0.
\]

Combining these results with the derivative of equation (5) with respect to \( \delta \) characterizes the condition for aggregate sales \( Q(c, \delta) \) to increase with price discrimination if:
\[ Q_\delta(c, \delta) = q_w \cdot q_s \cdot \frac{L^w \eta^w}{\pi_{pp} + \pi_{pp}} \cdot (\sigma^w - \sigma^s) > 0, \]  

(10)
i.e., when inverse demands are both convex and more so in the weak market, \( \sigma^w \geq \sigma^s \geq 0 \). This is true for log-concave demands and consistent with the constant-curvature demand functions of Bulow and Pfleiderer (1983) evaluated in Section 3. The result is important because for 3DPD to increase welfare a total output increase is necessary (Aguirre et al., 2010; Holmes, 1989; Schmalensee, 1981).

This setup with constant retail costs has been used for decades to find conditions where overall demand and welfare increase with 3DPD. It is however unlikely that a wholesaler with market power does not revise its pricing when aggregate sales \( Q(c, \delta) \) vary with retail price discrimination, i.e., unless \( Q_\delta(c, \delta) = 0 \). Simple inspection of equation (10) indicates that this is the case when the curvature of demand is identical across local markets, \( \sigma^w = \sigma^s \), which includes the case of linear demands of Schmalensee (1981, § II).

The optimal wholesale monopoly pricing solves:

\[ c(\delta) \in \arg\max_c \hat{\pi}(c, \delta) = (c - \zeta)Q(c, \delta), \]  

(11)which requires:

\[ \hat{L}(c, \zeta) = \frac{c - \zeta}{c} = \frac{1}{\hat{\eta}(c, \delta)}, \]  

(12)where \( \hat{\eta} \equiv -c \cdot Q_c(c, \delta)/Q(c, \delta) \) is the wholesale demand elasticity. We obtain the optimal wholesale price response to retail price discrimination by totally differentiating the upstream industry analog of equation 2, the wholesaler’s profit maximization condition:

\[ \frac{dc}{d\delta} = -\frac{\hat{\pi}_{c\delta}}{\hat{\pi}_{cc}}, \]  

(13)which equals zero only if the profit function is additively separable in \( c \) and \( \delta \). This occurs if all local demands have the same curvature and aggregate demand is additively separable in \( c \) and \( \delta \) so that \( \hat{\pi}_{c\delta} = Q_\delta(c, \delta) + (c - \zeta)Q_{c\delta}(c, \delta) = 0 \). Then \( Q(c, \delta = 0) = Q(c, \delta > 0) \) and the optimal wholesale price with retail discrimination is the same than with uniform retail pricing, \( c = c^u \).

Figure 1 explores the interaction between \( c \) and \( \delta \). The continuous blue line \( Q(c^u, \delta = 0) \) is the total wholesale demand when the retailer does not engage in price discrimination. This coincides with the horizontal summation of local retail marginal revenues with uniform pricing, \( \sum MR^u(p^u) \). Equating the corresponding marginal revenue function, dashed blue line \( \sum \hat{MR}(p^u) \), with the wholesaler’s constant marginal cost of production \( \zeta \) at (A) determines the optimal total sales \( Q_0 \) and wholesale price \( c^u \) at (B) with downstream uniform pricing. The figure depicts the case of welfare enhancing 3DPD where the overall output increases with \( \delta \), thus shifting demand to the thicker red \( Q(c^u, \delta > 0) \) line. Wholesaler’s optimal price \( c^u \) remains unchanged at (C) with this increased demand only if the corresponding marginal revenue function crosses marginal cost \( \zeta \) precisely at (D), with total sales \( Q_1 \). However, the increase in total sales depends on the position of each local demand and their relative curvatures \( \sigma^w \) and \( \sigma^s \) at \( p^w \) and \( p^s \), respectively.
Depending on demand specification, the resulting aggregate demand shifts to the right and may become steeper or flatter than the aggregate wholesale demand with uniform pricing. Ultimately, wholesale price $c$ increases with retail discrimination $\delta$ if the price elasticity of the new aggregate demand $Q(c^u, \delta > 0)$ evaluated at $c^u$ (C) is smaller than at the uniform pricing solution (B). If retail price discrimination makes wholesale demand less elastic, the upstream monopolist’s original markup is too low to maximize profits and wholesale price with retail price discrimination $c$ will increase relative to $c^u$. The opposite occurs if price discrimination turns wholesale demand more elastic than under uniform retail pricing.

Intuitively, an increase in wholesale demand following the retailer’s decision to price discriminate should trigger an increase in wholesale price. In order to ensure that $dc/d\delta \geq 0$ we only need to slightly restrict the behavior of wholesale demand $Q(c, \delta)$ and assume that the wholesaler’s profit function is log-supermodular in $c$ and $\delta$. This requires $\hat{\pi} \cdot \hat{\pi}_c \delta - \hat{\pi}_c \cdot \hat{\pi}_\delta \geq 0$ to ensure a monotone optimal wholesale price response \[13\] to retail price discrimination.\footnote{See Topkis (1998, §2.6.4) for a formal definition. Applications of log-supermodularity include auction theory (Milgrom and Weber 1982); international trade (Costinot 2009); monotone comparative statics (Milgrom and Shannon 1994); its preservation under uncertainty (Jewitt 1987); and in multivariate environments (Athey 2002).} Since log-supermodularity is preserved by multiplication, log-supermodularity of wholesaler’s profit function \[11\] just requires that total wholesale demand is also log-supermodular:

\[ \frac{\partial^2 \ln Q(c, \delta)}{\partial c \partial \delta} = -\frac{Q_c c \delta - Q_c c \cdot Q_\delta}{Q^2} \geq 0. \]
Notice that this condition is equivalent to the wholesale demand elasticity being non-increasing in \(\delta\), i.e., \(\partial \hat{\eta}(c, \delta)/\partial \delta \leq 0\). Furthermore, Topkis (1998, §2.6.4) shows that this demand elasticity being non-increasing for \(\delta \in \mathbb{R}\) implies that \(Q(c, \delta)\) is necessarily log-supermodular.\(^6\)

Therefore, if wholesale demand is less elastic at \(c_u\) with retail price discrimination than under retail uniform pricing, it is not reasonable to assume a constant wholesale price. Indeed, with wholesale constant returns to scale \(\zeta\) is constant and it follows from the wholesaler’s profit maximization condition (12) that retail price discrimination will induce an increase in wholesale price. This retail cost increase reduces retail profits (6) and its incidence on consumers depends on whether they belong to the weak or strong market, as the pass-through ratio is directly proportional to the curvature of demand. See equation (3). We need to determine next if this increase in retail costs increases or reduces the optimal price differential across local markets with price discrimination.

Continuing with the partial equilibrium approach, notice that \(\sigma^w \geq \sigma^s \geq 0\) suffices for 3DPD to increase sales and welfare as long as the wholesale price is constant, \(Q\delta(c^u, \delta) > 0\). But it then follows that a small increase in retail costs reduces the retail price spread:

\[
\frac{dp^s}{dc} - \frac{dp^w}{dc} = \frac{\sigma^s - \sigma^w}{(2 - \sigma^s)(2 - \sigma^w)} \leq 0.
\]

This result implies that if wholesalers enjoy some market power, retailers will not engage in 3DPD as aggressively as the observed differences in local demands suggest they should. This new result, previously neglected in the theoretical literature, as the important empirical implication that ignoring the ability of wholesalers to adjust their prices when retailers engage in price discrimination overestimates the profits derived by segmenting local markets. Alternatively, ignoring wholesale market power when evaluating the optimality of retail 3DPD overestimates the equilibrium managerial decision-making costs needed to rationalize the near-uniform pricing across markets observed in the data (Bloom and Van Reenen 2007; Della Vigna and Gentzkow 2017).

Our final result indicates that if the wholesaler has market power, the relative curvature condition (10) no longer suffices for aggregate sales and welfare to increase with 3DPD. To illustrate how several partial equilibrium results may lead to this conclusion we make use of Figure 2. First, because the wholesaler’s upward pricing revision, the aggregate demand increase due to price discrimination is lower than expected with a constant wholesale price as shown in equation (15). Thus, demand only shifts to the solid magenta line \(Q(c^*, \delta > 0)\) with \(c^* > c^u\). If wholesale price were to remain constant at \(c^u\) the new equilibrium at (E) would increase aggregate sales only from \(Q_0\) to \(Q_2\). As before, if wholesale demand elasticity at (E) is lower than at (B), the optimal wholesale price with 3DPD \(c^*\) exceeds the level \(c^u\) with retail uniform pricing. Notice however that if demand elasticity at (E) is sufficiently low relative to (B), it might be optimal for the wholesaler to increase its price so much to induce an overall reduction of sales \(Q_3 < Q_0\) at (F). Thus, the

\(^6\) Our problem is similar to Milgrom and Roberts (1990, §4(2)), who study conditions for own-price demand elasticity to be non-increasing in substitute product prices.
relative curvature of local demand condition \( \frac{\partial Q(c^u)}{\partial \delta} > 0 \) alone is no longer sufficient to ensure that 3DPD increases aggregate sales and enhances welfare.

In summary, our simple theoretical model predicts the following equilibrium effects of upstream market power:

1. If consumer demand is log-supermodular in \( c \) and \( \delta \), the upstream firm will increase its price when the downstream firm price discriminates; i.e., \( \frac{dc}{d\delta} \geq 0 \). Otherwise, downstream price discrimination induces the upstream firm to decrease its wholesale price.

2. The upward wholesale price revision always decreases downstream profits.

3. The reduction in profitability of retail 3DPD reduces price differentials across local markets: if price discrimination increases total sales, \( \frac{\partial Q(c^u)}{\partial \delta} > 0 \), and demand is log-supermodular, \( \frac{dc}{d\delta} \geq 0 \), the upstream firm’s increase in wholesale price reduces the dispersion of retail prices, \( \delta \downarrow 0 \).

4. The relative local demand curvature condition is no longer sufficient to ensure that 3DPD increases aggregate sales and thus, retail price discrimination may reduce overall aggregate sales.

Retail and wholesale price setting influence each other. The nature of upstream and downstream competition and demand and cost features determine the size of the aggregate demand
increase (or decrease), the wholesale price increase, the retail price spread, and the magnitude of managerial choice-making cost overestimation by misspecifying the upstream industry pricing behavior. The following two sections address the importance of wholesalers market power in reducing the incentive of retailers engaging in 3DPD, by how much wholesale price increases in response to retail price discrimination, and if aggregate sales fail to increase. Section 3 uses an analytical model where local demands are log-concave and have a constant curvature. Moving beyond the single-product case, Section 4 evaluates this diminished profitability effect using a much more flexible demand specification, not necessarily log-concave and certainly with a non-constant curvature, in the context of the Pennsylvania liquor market where a single monopoly retail state agency purchases from distillers with market power and enforces a uniform pricing policy.

3 Local Demand Convexity and Retail Price Discrimination

The partial equilibrium analysis of Aguirre et al. (2010) shows that output and welfare can increase when 3DPD allows for local market specific prices that are not too different from retail uniform pricing. In the neighborhood of the original uniform pricing it is possible to assume that changes in the curvature of local demands are negligible and thus easily characterize how 3DPD increases total output and welfare as a function of the curvature of local demands and not their change. The numerical analysis of this section assumes a constant curvature of local demands to address global solutions and to allow for larger deviations from uniform pricing equilibrium.

How much wholesalers revise their pricing when retailers decide to engage in 3DPD critically depends on the relative convexity of local inverse demands. The wholesaler’s demand increases with the difference \( \sigma^w - \sigma^s \geq 0 \). The more convex demand is in the weak market relative to the strong one. This section evaluates numerically how the wholesaler revises its pricing and reduces the retailer’s profits of engaging in 3DPD as a function of local demand curvature heterogeneity. To do so we follow Bulow and Pfleiderer (1983) and specify demands with constant inverse curvature:

\[
p = a - bq^{1-\sigma}, \quad \text{for } 0 < \sigma < 1.
\]

We characterize the equilibrium using this Bulow-Pfleiderer’s demand function when a monopolist retailer engages in 3DPD in two local markets or if instead it applies a uniform price across them, a result that is optimal only when both demand curvatures are identical. We address two scenarios: in the first, the retailer naïvely believes that upstream wholesale price remains invariant to the retailers’ pricing strategy choice while in the second, the retailer fully anticipates that the wholesale price will be revised to maximize monopoly profits.

3.1 Uniform Retail Pricing

A retail monopolist with constant marginal cost \( c \in [0, a] \) maximizes:
π = (p - c)q = (a - c)q - bq^{2-σ}, \hspace{1cm} (17)

with the following uniform retail pricing solution for any given wholesale price c:

\begin{align*}
q^u(c, σ) &= \left[ \frac{a - c}{b(2 - σ)} \right] \frac{1}{1 - \sigma}, \hspace{1cm} (18a) \\
p^u(c, σ) &= \frac{(1 - \sigma)a + c}{2 - \sigma} \geq 0, \hspace{1cm} \text{for } σ \leq 1 + \frac{c}{a}. \hspace{1cm} (18b)
\end{align*}

These expressions solve the expected downstream market solutions in the weak and strong markets, separately for $σ = σ^u$ and $σ = σ^s$, when a naïve retailer believes that the monopolist’s wholesale price $c$ will not change with the retailer’s decision to price discriminate.

When the retail monopolist does not price discriminate, the wholesale monopolist with constant marginal cost $ζ \in [0, c]$ maximizes:

\begin{equation}
\hat{π} = (c - ζ)q^u(c, σ) = (c - ζ) \left[ \frac{a - c}{b(2 - σ)} \right] \frac{1}{1 - \sigma}, \hspace{1cm} (19)
\end{equation}

at the following uniform price:

\begin{equation}
c^u(σ) = ζ + \frac{q^u(c, σ)}{q^u(c, σ)} = \frac{(1 - σ)a + ζ}{2 - σ} \geq 0, \hspace{1cm} \text{for } σ \leq 1 + \frac{ζ}{a}. \hspace{1cm} (20)
\end{equation}

After substituting $c^u(σ)$ into (18a)–(18b), the industry output and retail price equilibrium once the wholesaler re-optimizes pricing, are:

\begin{align*}
q^u(σ) &= q^u(c^u(σ), σ) = \left[ \frac{a - ζ}{b(2 - σ)^2} \right] \frac{1}{1 - σ}, \hspace{1cm} (21a) \\
p^u(σ) &= p^u(c^u(σ), σ) = \frac{(1 - σ)(3 - σ)a + ζ}{(2 - σ)^2}. \hspace{1cm} (21b)
\end{align*}

Finally, the profits of the wholesaler with uniform retail pricing are:

\begin{equation}
\hat{π}^u(σ) = \left[ \frac{(1 - σ)(a - ζ)}{2 - σ} \right] \left[ \frac{a - ζ}{b(2 - σ)^2} \right] \frac{1}{1 - σ}. \hspace{1cm} (22)
\end{equation}

As for the retailer’s profits with uniform pricing, it is worth distinguishing two expressions. First, $π^u(c, σ)$, represents the case where $c$ is either the competitive wholesale price or the optimal monopoly wholesale price when the retailer prices uniformly and is naïve regarding the pricing

---

7 Appendix A further addresses how demand curvature affects retail and wholesale profits when the retailer makes use of uniform pricing.
behavior of the wholesaler. Second, $\pi^u(\sigma)$, fully anticipates double marginalization and the wholesale price change induced by uniform retail pricing, i.e., after substituting $c = c^u(\sigma)$:

$$
\pi^u(c, \sigma) = \left[ \frac{(1 - \sigma)(a - c)}{2 - \sigma} \right] \left[ \frac{a - c}{b(2 - \sigma)} \right] \frac{1}{1 - \sigma},
$$

(23a)

$$
\pi^u(\sigma) = \pi^u(c^u(\sigma), \sigma) = \left[ \frac{(1 - \sigma)(a - \zeta)}{(2 - \sigma)^2} \right] \left[ \frac{a - \zeta}{b(2 - \sigma)^2} \right] \frac{1}{1 - \sigma}.
$$

(23b)

### 3.2 Retail Price Discrimination

To address how the equilibrium changes when the retailer engages in 3DPD, we assume that there is one weak market with inverse demand $p^w = a^w - b^w(q^w)^{1-\sigma^w}$ and a strong market with inverse demand $p^s = a^s - b^s(q^s)^{1-\sigma^s}$, so that the total wholesale monopolist’s demand when charging $c$ is:

$$
\tilde{Q}(c, \sigma^w, \sigma^s) = q^u(c, \sigma^w) + q^u(c, \sigma^s).
$$

(24)

In instead, the retail monopolist does not discriminate, which is optimal when $\sigma^w = \sigma^s$, the wholesale demand is:

$$
Q(c, \sigma^w, \sigma^s) = 2q^u(c, \sigma^w).
$$

(25)

To determine the optimal wholesale price $c^*$ when the retailer price discriminates, the wholesaler maximizes the following average profits per market:

$$
\hat{\pi}^*(c, \sigma^w, \sigma^s) = (c - \zeta)\tilde{Q}(c, \sigma^w, \sigma^s) = (c - \zeta) [q^u(c, \sigma^w) + q^u(c, \sigma^s)],
$$

(26)

which requires:

$$
[q^u(c, \sigma^w) + (c - \zeta)q^w(c, \sigma^w)] + [q^u(c, \sigma^s) + (c - \zeta)q^s(c, \sigma^s)] = 0.
$$

(27)

After replacing each solution $q^i(c, \sigma^i)$ for $i = \{w, s\}$ from (18a) and its derivative $q^i_x(c, \sigma^i)$, the optimal wholesale price with retail 3DPD is such that $c^u(\sigma^w) \leq c^*(\sigma^w, \sigma^s) \leq c^u(\sigma^s)$, see Appendix C and solves (numerically) the following condition:

$$
c^*(\sigma^w, \sigma^s) - \zeta = \left[ \frac{a^w - c^*}{b^w(2 - \sigma^w)} \right] \frac{1}{\sigma^w} + \left[ \frac{a^s - c^*}{b^s(2 - \sigma^s)} \right] \frac{1}{\sigma^s}.
$$

(28)
The retailer’s profits from 3DPD are:

$$\pi^*(c, \sigma^w, \sigma^s) = [p^u(c, \sigma^w) - c] q^u(c, \sigma^w) + [p^u(c, \sigma^s) - c] q^u(c, \sigma^s).$$  \hspace{1cm} (29)$$

This expression evaluated at $c = c^*$, correctly anticipates wholesalers with market power revising their pricing when retailers choose 3DPD instead of uniform pricing. The wholesale price revision leads to higher retailer costs, $c^* > c$, when wholesale demand elasticity is non-increasing in $\delta = p^s - p^w$. It must then be the case that retail profitability of 3DPD is also lower i.e., $\pi^*(c^*, \sigma^w, \sigma^s) \leq \pi^*(c, \sigma^w, \sigma^s)$. These lower retail profits reduce the incentive to engage in retail 3DPD and, as predicted in equation (15), the price differential between weak and strong markets is also smaller when the retailer anticipates the wholesale price change than when we wrongly assume to be constant, i.e., $p^u(c^*, \sigma^s) - p^u(c^*, \sigma^w) \leq p^u(c, \sigma^s) - p^u(c, \sigma^w)$.

Figure 3 summarizes the difference of the model predictions when wholesale price $c$ is assumed constant and when wholesalers revise it to its optimal level $c^*$. The analysis assumes a wholesale zero marginal cost and a retail choice between pricing two local stores separately or charging the same retail price across stores. Since price discrimination increases overall output when $c$ is constant and $\sigma^w \geq \sigma^s \geq 0$ we write $\gamma = \sigma^s/\sigma^w \in [0, 1]$ to evaluate how demand curvature heterogeneity across local markets affects the qualitative predictions of the model.

The dashed line in Figure 3 (a) plots the ratio of the average sales per market with price discrimination (24) and sales per market under uniform pricing (18a). This scale-free ratio that we label “Naïve Sales Ratio” is evaluated at $c$ and $\gamma = \sigma^w$. It shows that, in the spirit of equation (10), output is always expected to increase with 3DPD as long as $\sigma^w \geq \sigma^s \geq 0$ and wholesale price remains constant. Furthermore, the increase in aggregate output is larger the more pronounced the difference in demand curvature across weak and strong local markets is. But the actual output increase, the solid line in Figure 3 (a) is much smaller and even negative once we account for the wholesale price revision of the upstream monopolist to $c^*$.

Figure 3 (b) presents this “Actual Sales Ratio” for different values of $\gamma$, i.e., for varying degrees of demand heterogeneity across locations. The panel shows a result that is novel in the price discrimination literature. If the wholesale price does not remain constant, the increase in sales per market may be rather limited and retail 3DPD might even lead to an overall sales reduction even if the curvature of demand is larger in the weak than in the strong market. This is particularly true for large values of $\sigma^w$, when the curvatures of the weak and strong market are sufficiently different. The reason, to be discussed below, is that the wholesale price increases the most when local demands are very different, thus triggering a reduction of total sales. The key conclusion of this analysis is the theoretical predictions of welfare increases associated to 3DPD via overall market sales expansion, from Robinson (1933) to Aguirre et al. (2010), rely heavily on the assumption of a constant marginal cost and only apply when the wholesale industry is perfectly competitive.
Figure 3: Uniform Pricing vs. Two Local Markets

(a) 3DPD to Uniform Sales Ratios, ($\gamma = 0.5$)

(b) Actual Sales Ratio
\[ \frac{Q(c^*, \sigma_s, \sigma_w, \mu)}{Q(c^u, \sigma_s, \sigma_w)} \]

(c) Wholesale Price $c^*$

(d) Price Ratio
\[ \frac{p^u(c^*, \sigma_s, \sigma_w)}{p^u(c^u, \sigma_s, \sigma_w)} \]

(e) 3DPD to Uniform Profit Ratios, ($\gamma = 0.5$)

(f) Profit Ratio
\[ \frac{\pi^u(c^*, \sigma_s, \sigma_w, \mu)}{\pi^u(c^u, \sigma_s, \sigma_w, \mu)} \]
In Figure 3 (c) we turn to the optimal wholesale pricing decision under retail price discrimination. Uniform pricing is optimal when all local markets have the same demand, $\gamma = 1$. This case determines the value $c$ for each $\sigma^w$ as shown in equation (20) and is represented by the (lowest) black monotone decreasing function in Panel (c). This is the value of $c$ that the retailer mistakenly assumes constant when engaging in 3DPD ignoring the wholesaler’s market power. The optimal wholesale price $c^*$ depends on the difference in curvatures $\sigma^w$ and $\sigma^s$ as shown in equation (28). Figure 3 (c) shows that $c^*(\sigma^w, \sigma^s) > c^u(\sigma^w)$ always, and that optimal wholesale price with retail price discrimination $c^*$ increases more relative the wholesale price with uniform retail pricing $c$ the more heterogeneous the curvatures of inverse demands in the weak and strong markets are.

Next, Figure 3 (d) shows that the upward revision in retail costs triggered by the expansion of the downstream market (and upstream demand) translates into smaller optimal price differences across local markets than acknowledged by the existing theoretical literature and current empirical studies: the spread of prices across weak and strong markets is now smaller as postulated in equation (3). Panel (h) shows the (scale free) price difference ratio $[p^u(c^*, \sigma^s) - p^u(c^u, \sigma^w)]/[p^u(c^u, \sigma^s) - p^u(c^u, \sigma^w)]$. Notice that actual price spread is a fraction of what theoretical models assuming a constant wholesale price will predict. This fraction is smaller the more convex inverse demands are and the more heterogeneous demands in weak and strong markets are.

How does the wholesaler’s price revision affect the profitability of retail 3DPD? Figure 3(e) presents two ratios, $\pi^*(c^*, \sigma^w, \sigma^s, \mu)/\pi^u(c^u, \sigma^w)$ and $\pi^*(c^u, \sigma^w, \sigma^s, \mu)/\pi^u(c^u, \sigma^w)$ evaluated for the particular case when the curvature of the inverse demand of the weak market is double that of the strong market case, $\gamma = 0.5$. Qualitative results hold for any other value of $\gamma \in [0, 1]$. The first ratio, represented by the continuous line, indicates that actual retail profits increase only a little relative to the profits of uniform retail pricing when the wholesaler revises its pricing in response to the retailer’s decision to price discriminate. The second ratio, represented by the dashed line, shows this same profit ratio when the retailer fails to anticipate the increase in retail costs due to the wholesale price revision when retailers engage in 3DPD. Thus, ignoring the ability of wholesalers with market power to revise their pricing may lead retailers (and practitioners) to overestimate the profitability of retailers engaging in 3DPD relative to uniform pricing.

To complete the analysis Figure 3(f) presents the profit ratio $\pi^*(c^*, \sigma^w, \sigma^s, \mu)/\pi^u(c^u, \sigma^w, \sigma^s, \mu)$ and how it varies with the heterogeneous convexity of local demands in order to evaluate whether retail profits are larger or smaller with price discrimination when wholesale price is revised upwards. Once we account for the wholesaler’s pricing revision, the expected retail profits of 3DPD are only a fraction of the naïve expected profits relative to those from optimal uniform pricing. The reduction in profits is more important the more convex local inverse demands are and the larger the difference in curvature across local demands. It should be noted that actual retail profits when $c = c^*$ can be substantially smaller (up to 60% to 70%) than their naïve counterparts when $c = c^u$. This raises an important empirical issue because a misspecified model wrongly assuming that the wholesale market is competitive may lead to a substantial overestimation of the lost rents for retailers not engaging in 3DPD.
To summarize, when wholesalers have market power, it is no longer valid to assume that a retailer’s decision to price discriminate does not affect the wholesale price. Indeed, wholesale price increases if local demand curvatures are such that $3DPD$ would increase overall sales with a constant wholesale price. Total sales, however, might end up being lower under $3DPD$ particularly for high values of $\sigma$, i.e., when the pass-through rates of log-concave demands is nearly complete. Retailers still profit from $3DPD$ but the profitability of this pricing strategy is substantially lower than when wholesale price $c$ remains constant. Consequently, the price spread between weak and strong markets under $3DPD$ is also significantly reduced.$^8$

---

$^8$ Table C.1 in the Appendix quantifies the expected effect of engaging in retail $3DPD$ on all these economic variables by integrating out over $\sigma^w$ and for different values of $\gamma$ accounting for heterogeneity of demand across local markets.
4 The Pennsylvania Liquor Market

4.1 Data Overview

Our data, obtained under the Pennsylvania Right-to-Know Law, contain daily information on quantities sold and gross receipts at the UPC and store level for all spirits carried by the PLCB during 2003 and 2004. We aggregate the daily data to the monthly sales pricing periods, resulting in 22 periods for our sample. The PLCB also provided the wholesale cost of each product, which is constant across stores, but varies over time reflecting the temporary or permanent price changes.

4.2 Data: Quantities Sold, Prices, and Characteristics of Spirits

Each store carries a vast variety of products among which we focus on popular 375 ml, 750 ml, and 1.75 L spirits products, representing 64.1% of total spirit sales measured in bottles and 70.1% of total spirit sales by revenue. The resulting sample contains 312 products across the two-year sample. We see little variation in the product set across stores indicating that consumers have equal access to the products we consider in the sample.

We classify products into six categories: brandy, cordials, gin, rum, vodka, and whiskey. For each product, the PLCB provided its alcohol content measured as proof (100 proof corresponds to 50% alcohol content by volume) and whether or not the product is imported or contains flavorings. Vodkas and whiskeys have significantly larger market shares (32.1% and 24%, respectively) than rum (16.3%), cordials (13.6%), or brandy (7.3%), even though cordials is one of the top categories in terms of number of products. Flavored spirits, which represent 16.3% of products, are primarily cordials and brandies and, to a lesser extent, rums and vodkas. Most whiskeys and cordials are imported while other spirits are predominantly domestically produced. There is significant variation in proof across product categories: the average across all products is 75.33, but it ranges from 55.82 for cordials to 83.42 for gins. We also obtained a product score rating products in each spirit category as a measure of within-category product quality from Proof66.com, a spirits ratings aggregator.

To report results and evaluate the diverse demand across demographic groups we characterize spirits as expensive when their simple averaged price exceeds the mean price of all spirits of the same type and bottle size. Expensive products are less likely to be flavored or domestically produced and have higher proof, but consumers purchase them nearly as frequently as cheaper ones. The 750 ml bottle is the most popular size in terms of unit sales and product variety, accounting for 50.3% of unit sales and 54.5% of available spirits products, followed by the 1.75 L bottle with a

---

9 Many products are available to consumers but are seldom purchased. The 375 ml, 750 ml, and 1.75 L bottle sizes account for 80.9% of total spirit sales by volume and 91.6% of total spirit sales by revenue. Within these bottle sizes, we further focus on popular products that account for 80% of bottle sales in each spirit type-bottle size combination. We also drop tequilas, as there were few products and these products amounted to only 1.6% of total liquor bottle sales. In total, these restrictions allow us to drop a total of 1,313 products from our sample.

10 The median store carries 98% of the top 100 and 82% of the top 1,000 products.
share of 34.5% of unit sales and 30.1% of products. The smallest bottles we consider, those in the 375 ml format, account for the remaining 15.2% of units sold and 15.4% of products.

These patterns in market shares reflect in part the product sets offered by distillers as not all brands are available in all bottle sizes. For instance, our final sample consists of 198 brands (e.g., Captain Morgan). Of these, 88 are available only in the 750 ml bottle size and one and 31 only in the 375 ml and 1.75 L sizes, respectively. The PLCB carries at least two bottle sizes for the remaining 78 brands (e.g., Diageo sold Captain Morgan in 375 ml, 750 ml, and 1.75 L sizes).

4.2.1 The Upstream Distillers

During our sample period, 34 firms compete in the spirits market. The market leader, Diageo, accounts for 22% of total unit sales and 25% of revenue, while a large set of small fringe producers (29) account for 42% and 46% of total quantity sold and revenue, respectively. Nineteen of these firms operate product portfolios of less than five products, and seven are single product firms. Table 1 documents that while large firms such as Diageo and Bacardi operate extensive product portfolios, there is substantial heterogeneity across firms in their product offerings. For example, Diageo has a relatively balanced portfolio where rums, vodkas, and whiskies generate approximately 21%, 31%, and 25% of revenue, respectively. In contrast, Bacardi operates a more concentrated portfolio as 71% of its revenue comes from its rum products compared to 19% from its whiskey products. Among the larger competitors, only the Pennsylvania-based firm Jacquin sells brandies, where it faces only seven small competitors and generates 22% of its revenue. With a presence in all bottle sizes and spirit types, the company’s portfolio focuses exclusively on cheap products. Table 1 furthermore documents that a significant number of competitors are present in all product categories. The variation in product portfolios translates into variation in concentration across categories, with Herfindahl-Hirschman indices ranging from 1.023 for cordials to 3.087 for rums (951 for spirits in total). This, combined with the observed degree of product differentiation, motivates our characterization of the distillery market as oligopolistic.

4.3 Theory

In this section we outline the theoretical model which underlies the empirical analysis of this section. Our environment consists of upstream distillers which choose prices simultaneously in each period $t$. Given these prices, a downstream monopoly chooses whether to price discriminate by setting store-level prices; i.e., whether or not to $3DPD$. If it chooses $3DPD$, it incurs a fixed fee $F \geq 0$ and chooses the vector store-level prices $\{p_{jl}\}$. If not, it chooses a uniform price $\{p_{jt}\}$. For now, we consider the limiting case where $F = 0$ so the downstream monopolist always chooses to $3DPD$. In the following empirical analysis, we compare the pay-offs between uniform and $3DPD$ to test the predictions from the simple, but also more general, theoretical framework presented in Section 2.

Given retail prices, consumers choose the product which maximize period $t$ utility. Since we do not observe evidence of dynamic decisions among either firms or consumers in our data (e.g.,
Table 1: Upstream Product Portfolios

<table>
<thead>
<tr>
<th></th>
<th>Diageo</th>
<th>Bacardi</th>
<th>Beam</th>
<th>Jacquin</th>
<th>Sazerac</th>
<th>Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>By Spirit Type:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brandy</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>21.57</td>
<td>0.00</td>
<td>8</td>
</tr>
<tr>
<td>Cordials</td>
<td>11.22</td>
<td>1.98</td>
<td>16.22</td>
<td>9.77</td>
<td>13.29</td>
<td>18</td>
</tr>
<tr>
<td>Gin</td>
<td>11.67</td>
<td>8.26</td>
<td>5.14</td>
<td>1.77</td>
<td>3.62</td>
<td>10</td>
</tr>
<tr>
<td>Rum</td>
<td>21.27</td>
<td>70.89</td>
<td>4.06</td>
<td>22.12</td>
<td>0.00</td>
<td>10</td>
</tr>
<tr>
<td>Vodka</td>
<td>30.99</td>
<td>0.00</td>
<td>10.4</td>
<td>43.26</td>
<td>63.9</td>
<td>14</td>
</tr>
<tr>
<td>Whiskey</td>
<td>24.85</td>
<td>18.88</td>
<td>64.18</td>
<td>1.52</td>
<td>19.18</td>
<td>20</td>
</tr>
<tr>
<td><strong>By Price:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cheap</td>
<td>19.60</td>
<td>34.20</td>
<td>52.61</td>
<td>100.00</td>
<td>85.30</td>
<td>25</td>
</tr>
<tr>
<td>Expensive</td>
<td>80.40</td>
<td>65.80</td>
<td>47.39</td>
<td>0.00</td>
<td>14.70</td>
<td>25</td>
</tr>
<tr>
<td><strong>By Bottle Size:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>375 ml</td>
<td>5.82</td>
<td>5.90</td>
<td>2.93</td>
<td>10.24</td>
<td>15.07</td>
<td>18</td>
</tr>
<tr>
<td>750 ml</td>
<td>54.67</td>
<td>50.59</td>
<td>38.13</td>
<td>27.05</td>
<td>20.53</td>
<td>31</td>
</tr>
<tr>
<td>1.75 L</td>
<td>39.51</td>
<td>43.51</td>
<td>58.95</td>
<td>62.71</td>
<td>64.59</td>
<td>25</td>
</tr>
<tr>
<td>All Products</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>34</td>
</tr>
</tbody>
</table>

Notes: Table displays firms’ revenue share by spirit type, price, and bottle size. “Firms” is the total number of firms with at least one product in the given category.

stockpiling), we do not consider the possibility that agents maximize pay-offs across periods. In the subgame perfect Nash equilibrium, however, period $t$ upstream firms do correctly anticipate the monopolist’s downstream pricing decisions as well as and purchase decisions of consumers.

### 4.3.1 Consumer Demand

We follow the large literature on discrete-choice demand system estimation using aggregate market share data [Berry, 1994; Berry, Levinsohn and Pakes, 1995; Nevo, 2001] to model demand for spirits as a function of product characteristics and prices. By mapping the distribution of consumer demographics into preferences, the model enables us to estimate realistic substitution patterns between products while accounting for the heterogeneity in preferences exhibited in Table 2. We assume that consumer $i$ in market $l$ in period $t$ obtains the following indirect utility from consuming a bottle of spirit $j \in J_{lt}$

$$u_{ijlt} = V_{ijlt} + \epsilon_{ijlt} = x_j \beta_i + \alpha_i p_j^r + H_t \gamma + \xi_{ijlt} + \epsilon_{ijlt},$$

where $i = 1, \ldots, M_{lt}; \quad j = 1, \ldots, J_{lt}; \quad l = 1, \ldots, L; \quad t = 1, \ldots, T.$

---

11 In the absence of individual purchase information we opt to treat bottles of different sizes of the same spirit as different products with identical observable characteristics other than size. It is likely that firms use bottle size as a quantity discount and therefore they set product prices jointly, solving a second degree price discrimination problem as consumers self-select into different bottle sizes depending on their willingness to pay. Modeling such second degree price discrimination formally requires accounting for informational asymmetries between distillers and consumers. Given the reasonable substitution patterns in the estimated model, we believe our approach is a good first approximation to the complex pricing problem of second degree price discrimination among multi-product firms in oligopoly competition, especially given that, to our knowledge, there does not exist a tractable theoretical model of nonlinear multi-product oligopoly pricing.
The $n \times 1$ vector of observed time-invariant product characteristics $x_j$ is identical in all markets $l$, though the availability of different spirits changes over time due to product introductions or removals. The $T \times 2$ matrix $H_t = [q_3 t \quad m 12_t]$ includes a summer dummy for periods in July, August, and September and a holiday dummy for periods $t$ coinciding with the holiday season from Thanksgiving to the end of the year. We denote the price of product $j$ at time $t$ by $p_{jt}^*$; it is the same across all markets $l$. We further allow utility to vary across products, markets, and time via the time and location-specific product valuations $\xi_{jlt}$, which are common knowledge to consumers, firms, and the PLCB but unobserved by the econometrician. Lastly, $\epsilon_{ijlt}$ denotes idiosyncratic unobserved preferences by consumer $i$ for product $j$ in market $l$ and period $t$, which we assume to be distributed Type-I extreme value across all available products $J_{lt}$.

We characterize consumer $i$ in market $l$ by a $d$-vector of observed demographic attributes, $D_{il}$ including age, education, income, and race. To allow for individual heterogeneity in purchase behavior and relax the restrictive substitution patterns inherent in the multinomial Logit, we model the distribution of consumer preferences over characteristics and prices as multivariate normal:

$$ \left( \frac{\alpha_i^*}{\beta_i^*} \right) = \left( \frac{\alpha}{\beta} \right) + \Pi D_{il} + \Sigma \nu_{il}, \quad \nu_{il} \sim N(0, I_{n+1}), $$

(31)

where $\nu_{il}$ captures mean-zero, unobserved preference shifters with a diagonal variance-covariance matrix $\Sigma$ (i.e., $\Sigma_{jk} = 0 \forall k \neq j$). $\Pi$ is a $(n + 1) \times d$ matrix of coefficients that measures the effect of observable individual attributes on the consumer valuation for spirit characteristics including price, allowing cross-price elasticities to vary differentially in markets with observed differences in demographics.

We make the common assumption that during period $t$, each consumer either selects one of the $J_{lt}$ spirits available in her market or chooses the outside option. We define the potential market, $M_{lt}$, to be the consumption of all alcoholic beverages off the premise of the seller, i.e., not in a restaurant or bar (“off-premise consumption”), during pricing period $t$. We calculate $M_{lt}$ as the number of drinking-age residents scaled by per-capita off-premise consumption, where we allocate the available annual per-capita consumption evenly across pricing periods according to the periods’ lengths. The outside option thus consists of closed-container beer or wine purchases denominated in 750 ml bottle-equivalents.

12 Nevo (2000) discusses limitations of the present discrete choice approach when individuals purchase several products or multiple bottles of the same product at the same time. If such consumer behavior were important, Hendel (1999) and Hendel and Nevo (2006) show that assuming single-unit purchases could underestimate price elasticities in the case of assortment decisions, but overstate own-price elasticities in the case of stockpiling. In Miravete et al. (2018a), we test for stockpiling using a similar dataset and find no evidence. Seim and Waldfogel (2013) present suggestive evidence that the PLCB’s demand does not respond disproportionately to price declines in areas where consumers have higher travel costs to the store and thus a higher incentive to buy larger quantities or assortments.

13 This definition of the potential market accounts for the total volume of alcoholic beverages but not for the different average ethanol contents of beer (4.5%), wine (12.9%), and spirits (37.7%) in our sample.

14 For example, according to Haughwout, Lavallee and Castle (2015), the average drinking-age Pennsylvanian consumed 96.2 liters of alcoholic beverages through off-premise channels in 2003, or 128.2 750 ml bottle equivalents.
Consumer utility-maximization connects the set of individual-specific attributes and the set of product characteristics as follows

\[ A_{jt}(x, p^r_t, \xi_t; \theta) = \{(D_{it}, \nu_{it}, \epsilon_{i-it}) | u_{ijlt} \geq u_{iklt} \ \forall k = 0, 1, \ldots, J_{lt}\}, \]

(32)

where we summarize all model parameters by \( \theta \). We follow the literature in decomposing the deterministic portion of the consumer’s indirect utility into a common part shared across consumers, \( \delta_{jtlt} \), and an idiosyncratic component, \( \mu_{ijlt} \), given by

\[
\begin{align*}
\delta_{jtlt} &= x_j \beta + \alpha p^r_{jt} + H_t \gamma + \xi_{jlt}, \\
\mu_{ijlt} &= \left( x_j p^r_{jt} \right) \left( \Pi D_{it} + \Sigma \nu_{it} \right).
\end{align*}
\]

(33a) \hspace{1cm} (33b)

In estimating the model, we integrate over the distribution of \( \epsilon_{i-it} \) analytically. The probability that consumer \( i \) purchases product \( j \) in market \( l \) in period \( t \) is then

\[
s_{ijlt} = \frac{\exp \left( \delta_{jtlt} + \mu_{ijlt} \right)}{1 + \sum_{k \in J_{lt}} \exp(\delta_{ktlt} + \mu_{iktlt})}.
\]

(34)

Deriving product \( j \)'s aggregate market share in each location requires integrating over the distributions of observable and unobservable consumer attributes \( D_{it} \) and \( \nu_{it} \), denoted by \( P_D(D_i) \) and \( P_\nu(\nu_i) \), respectively. The market share for product \( j \) in market \( l \) at time \( t \) is:

\[
s_{jlt} = \int_{D_i} \int_{\nu_l} s_{ijlt} dP_D(D_i) dP_\nu(\nu_i),
\]

(35)

which we evaluate using simulating techniques.

An advantage of a structural model is that it enables the researcher to assess equilibrium changes in welfare. At retail prices \( p^r \), the (expected) consumer surplus of consumer \( i \) in location \( l \) at period \( t \) is

\[
CS_{ilt}(p^r) = \frac{1}{\alpha_i^*} \times \sum_{j \in J_{lt}} \exp \left[ V_{ijlt}(p^r_t) \right] + C,
\]

(36)

where \( C \) is an unknown constant on integration reflecting the fact that the absolute level of consumer utility cannot be measured. We identify beneficiaries of the single markup policy by evaluating changes in consumer welfare via compensating variation, i.e., the amount of income necessary to keep individuals in a given market indifferent between any counterfactual set of prices \( p''^r \) and the current ones \( p^r \). Since consumer utility is quasi-linear, changes in retail prices generate no income.

---

2003 potential market for location \( l \) is then the number of drinking-age residents scaled by 128.2. To put this figure in perspective, beer accounts for approximately 90% of total consumption by volume so the average drinking-age Pennsylvanian consumed slightly less than five 375 ml bottles of beer per week, but only approximately thirteen 750 ml bottles of both wine and spirits annually. We follow a similar approach in constructing the potential market for 2004.
effects so the Marshallian demand is equivalent to Hicksian demand. As a result, therefore changes in consumer surplus (CS) are equivalent to compensating variation:

$$CV_{ilt}(p^r, p^{r'}) = \frac{1}{\alpha^*_i} \ln \left[ \frac{\sum_{j \in J_{lt}} \exp \left[ V_{ijlt}(p^{r'}) \right]}{\sum_{j \in J_{lt}} \exp \left[ V_{ijlt}(p^r) \right]} \right],$$

(37)

where $V_{ijlt}(\cdot)$ is given by (30). The mean compensating variation for agents living in location $l$ is

$$CV_l(p^r, p^{r'}) = \sum_t M_{lt} \int_{\nu_i} \int_{D_l} CV_{ilt}(p^r, p^{r'}) dP_{D_l} dP_{\nu_i}.$$  

(38)

Residents in location $l$ are thus on average better off under the current policy when $CV_l(p) > 0$, indicating that they require positive compensation to be unaffected by the new policy with retail prices $p'$.

4.3.2 Downstream Monopoly Pricing

Given period $t$ upstream wholesale prices $\{p^w_{jt}\}$ consumer demand, the downstream monopolist chooses the vector of store-level prices $\{p^r_{jlt}\}$ to maximize period $t$ profit:

$$\max_{p^r_{jlt}} \sum_{j \in J} \left[ (p^r_{jlt} - p^w_{jt}) \times s_{jlt}(p^r, x, \xi; \theta) \times M_{lt} \right].$$

(39)

Profit maximization in the downstream market implies the following period $t$, product $j$ first-order condition for market $l$:

$$s_{jlt}(p^r, x, \xi; \theta) + \sum_{m \in J} \left( p^r_{mlt} - p^w_{mlt} \right) \times \frac{\partial s_{mlt}}{\partial p^r_{jlt}} = 0.$$  

(40)

Equation (39) demonstrates a key feature of our problem as the upstream firms set statewide, product-level prices which our downstream retailer takes as its sole marginal cost but our retailer possesses the ability to charge product-level prices which vary by store. Thus, our model is similar to the case of the chain-store considered by Della Vigna and Gentzkow (2017).

4.3.3 Upstream Oligopoly Pricing

Given optimal downstream pricing and consumer choices, we now consider competition between distillers. A total of $F$ firms compete in the upstream market where each firm $f \in F$ produces a subset $J_{lt}^f$ of the $j = 1, \ldots, J_t$ products. We assume that in each period $t$, distillers set wholesale
prices vector of wholesale prices \( \{ p_{jt}^w \}_{j \in J^f} \) non-cooperatively as in a Bertrand-Nash differentiated products oligopoly to maximize period \( t \) profit

\[
\max_{p_{jt}^w} \sum_{j \in J^f} \left[ (p_{jt}^w - c_{jt}) \times \sum_{l=1}^{L} M_l s_{jlt}(p^r(p^w), x, \xi; \theta) \right],
\]

(41)

where \( c_{jt} \) denotes the marginal cost of product \( j \) in period \( t \). Define as \( s_{jt}(p^r, x, \xi; \theta) \) the statewide demand for product \( j \) in period \( t \), \( \sum_{l=1}^{L} M_l s_{jlt}(p^r, x, \xi; \theta) \). Profit maximization in the upstream market implies the following first-order condition for distiller \( f \)'s product \( j \), \( \forall j \in J^f \):

\[
s_{jt}(p^r(p^w), x, \xi; \theta) + \sum_{m \in J^f} \left( p_{mt}^w - c_{mt} \right) \times \frac{\partial s_{mt}}{\partial p_{jt}^w} = 0.
\]

(42)

The final term \( \frac{\partial s_{mt}}{\partial p_{jt}^w} \) is the response in product \( m \)'s quantity sold to a change in the wholesale price and, through the downstream pricing response, the retail price of product \( j \). Assuming a pure-strategy equilibrium in wholesale prices, the vector of profit-maximizing wholesale prices is

\[
p_{jt}^w = c_t + \left[ O_t^w \ast \Delta_t^w \right]^{-1} \times s_t(p^r(p^w), x, \xi; \theta),
\]

(43)

where \( O_t^w \) denotes the ownership matrix for the upstream firms with element \((j, m)\) equal to one if goods \( j \) and \( m \) are in \( J^f \) and firm \( f \) chooses these prices jointly. We define \( \Delta_t^w = -\Delta_t^d \Delta_t^p \) as a matrix that captures changes in demand due to changes in wholesale price,

\[
\Delta_t^w = - \begin{bmatrix} \frac{\partial s_{11}}{\partial p_{11}^w} & \cdots & \frac{\partial s_{1L}}{\partial p_{1L}^w} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_{J1}}{\partial p_{11}^w} & \cdots & \frac{\partial s_{JL}}{\partial p_{1L}^w} \end{bmatrix} \times \begin{bmatrix} \frac{dp_{11}^r}{dp_{11}^w} & \cdots & \frac{dp_{1L}^r}{dp_{1L}^w} \\ \vdots & \ddots & \vdots \\ \frac{dp_{J1}^r}{dp_{11}^w} & \cdots & \frac{dp_{JL}^r}{dp_{1L}^w} \end{bmatrix},
\]

(44)

where \( \Delta_t^d \) is the matrix of changes in quantity sold due to changes in retail price with element \((k, m)\) equal to \( \frac{\partial s_{km}}{\partial p_{mt}^w} \) and \( \Delta_t^p \) is the matrix of changes in retail price due to changes in wholesale price with element \((m, j)\) equal to \( \frac{dp_{rmt}}{dp_{jt}^w} \). Given all agents have perfect information, we assume the upstream firms correctly anticipate the downstream firm response matrix in each period \( t \) subgame perfect Nash equilibrium. In the estimation, however, the state’s regulation of alcohol sales commits it to applying a uniform ad valorem markup so this matrix simplifies significantly by eliminating off-diagonal terms so that \( \frac{dp_{rmt}}{dp_{jt}^w} \) is simply \( 1.30 \times 1.18 = 1.534 \) for all stores, reflecting the 30% uniform markup and the 18% liquor tax.

Given estimates of consumer demand, data on retail and wholesale prices, and the observed \( PLCB \) policy, we use this model of upstream behavior to recover product-level marginal costs via (43). These marginal cost estimates ultimately enable us to evaluate the equilibrium effects of downstream \( 3DPD \).
Table 2: Connecting Consumer Preferences and Demographics

<table>
<thead>
<tr>
<th></th>
<th>RISK</th>
<th>AGE</th>
<th>MINORITY</th>
<th>EDUCATION</th>
<th>INCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>By Spirit Type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRANDY</td>
<td>6.6</td>
<td>7.3</td>
<td>9.8</td>
<td>4.6</td>
<td>5.2</td>
</tr>
<tr>
<td>CORDIALS</td>
<td>14.7</td>
<td>12.9</td>
<td>13.2</td>
<td>13.4</td>
<td>15.3</td>
</tr>
<tr>
<td>GIN</td>
<td>5.3</td>
<td>8.0</td>
<td>8.1</td>
<td>6.4</td>
<td>4.7</td>
</tr>
<tr>
<td>RUM</td>
<td>18.4</td>
<td>15.4</td>
<td>17.5</td>
<td>14.2</td>
<td>16.6</td>
</tr>
<tr>
<td>VODKA</td>
<td>27.7</td>
<td>34.5</td>
<td>32.5</td>
<td>34.1</td>
<td>27.4</td>
</tr>
<tr>
<td>WHISKEY</td>
<td>27.3</td>
<td>21.9</td>
<td>18.8</td>
<td>27.4</td>
<td>30.9</td>
</tr>
<tr>
<td>By Price:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHEAP</td>
<td>56.1</td>
<td>50.7</td>
<td>55.0</td>
<td>49.1</td>
<td>54.8</td>
</tr>
<tr>
<td>EXPENSIVE</td>
<td>43.9</td>
<td>49.3</td>
<td>45.0</td>
<td>50.9</td>
<td>45.2</td>
</tr>
<tr>
<td>By Bottle Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>375 ml</td>
<td>15.1</td>
<td>15.7</td>
<td>20.1</td>
<td>10.6</td>
<td>10.5</td>
</tr>
<tr>
<td>750 ml</td>
<td>49.9</td>
<td>50.6</td>
<td>51.1</td>
<td>49.1</td>
<td>49.3</td>
</tr>
<tr>
<td>1.75 L</td>
<td>35.0</td>
<td>33.7</td>
<td>28.8</td>
<td>40.3</td>
<td>40.2</td>
</tr>
</tbody>
</table>

Notes: Table displays market shares based on bottles sold by product characteristic for markets in the bottom ("Low") and top ("High") quintiles of each demographic attribute. RISK denotes per capita ethanol consumption. Other variables defined as average age in the market (AGE); share of population with some college education (EDUCATION); share of non-white population (MINORITY); and share of households with income greater than $50,000 (INCOME). Figure 4 displays the spatial distribution of demographics. Source: 2000 Census of Population.

4.4 Demand Heterogeneity

We complete this descriptive analysis by documenting the heterogeneity of preferences for spirit types and product characteristics across demographic attributes. We compare unit-sales market shares for various product categories across demographic market groupings. For instance, the top quintile of markets by income garner a 38.6% share of 1.75 L bottles. Table 2 highlights important market share differences across demographics (columns) for different product categories (rows). These purely reflect differences in preferences since retail prices at a point in time are identical across the state and stores have similar product offerings.

The data indicate that minorities strongly favor brandy, gin, and 375 ml products, but not whiskey or 1.75 L products. In markets with high income and a highly educated population, vodka is far more popular than rum and brandy while consumers also buy spirits that are more expensive. Markets dominated by young, less educated, and lower income populations show a clear preference for cheap products. The popular 750 ml bottle amounts to almost exactly half of all bottle sales across demographic attributes, but between the 375 ml and 1.75 L sizes, higher-income markets clearly favor 1.75 L products. Finally, heavy drinkers, i.e., consumers in high-risk markets, prefer expensive and vodka products, but are unlikely to purchase 375 ml bottles, reflecting a positive correlation between per capita ethanol consumption and income.

Figure 4 documents a great deal of demographic diversity of Pennsylvania. Combined with Table 2 this suggests the possibility that local pricing (i.e., 3DPD) may lead to significant increases in profitability.
To summarize, the data are useful in investigating the effectiveness of 3DPD for three reasons. First, we observe all spirit sales in the state during the sample period which enables us to model an entire industry. Second, as in the case of Della Vigna and Gentzkow (2017), our product set is heterogeneous, varies little across retail stores, and product characteristics correlate systematically with consumer purchases. Unlike Della Vigna and Gentzkow (2017), however, our data enable us to incorporate substitution across products within a store as well as the strategic interaction between upstream firms and the downstream monopolist. Thus, we are able to quantify the equilibrium effects of 3DPD on not only on a diverse set of consumers and firms.

4.5 Estimation Results

We adapt the estimation approach of Nevo (2000) to the institutional features surrounding the price regulation of spirits in Pennsylvania. We follow a three-step estimation procedure that takes advantage of the fact that the PLCB charges the same retail price for a given product in all local markets. This allows us to identify separately the contribution to demand of demographic taste heterogeneity across the state at a point in time and the contribution of time varying shifters of demand that are common across demographic groups, including price. In the first-stage, we
use generalized method of moments techniques to estimate the determinants of deviations from a product’s mean utility \(\mu_{ijlt}\), controlling for market and product by pricing period fixed effects that absorb the mean effects of price, product characteristics, and seasonality. In the second stage, we employ instrumental variables techniques to project the estimated product by pricing period fixed effects onto price, seasonality, and product fixed effects, using contemporaneous prices in distant control states and input prices as instruments. Last, we project the estimated product fixed effects from the second stage onto time-invariant product characteristics. We refer the interested reader to our companion piece, Miravete et al. (2018b), which provides details of the estimation procedure for a more parsimonious version of the demand model that we estimate here.

Before discussing the estimation results, we turn to the variation in the data that allows identification of two key aspects of the demand model: heterogeneity in consumer preferences and consumer price sensitivity. We identify unobserved heterogeneity in consumer preferences – the random coefficients \(\Sigma\) in equation (31) – from correlation between a product’s market share and its characteristics relative to other more or less similar products; see Berry and Haile (2014). We construct two instruments similar to those used in Bresnahan, Stern and Trajtenberg (1997). First, we employ the number of products in the market that share product \(j\)’s characteristic. For example, to identify taste variation for brandies, we use the total number of competing brandies of the same bottle size in location \(l\) in period \(t\) as the instrument for a given brandy. Second, we use the root mean square distance in spirit product scores, as a measure of product quality, of product \(j\) and other products that share its characteristics. Thus, for the above brandy, this would be the average product score distance from other brandies available in market \(l\) at time \(t\). This instrument provides additional identifying power since it captures the differential effect on the market share of a high-quality brandy, say, as the product quality of the brandies that it competes against changes, e.g., over time.

Variation in consumer preferences due to demographic variation across markets – the demographic interactions \(\Pi\) in equation (31) – reflects correlation between the market shares of products with particular characteristics in a given store market and the demographics of the population served by each store. A key feature aiding identification of \(\Pi\) is the uniform retail price for each product across markets, facilitating the linking of purchases to demographic variation in preferences alone. Following Waldogel (2003), we interact the above two instruments with the prevalence of a given demographic attribute in each market. For example, we would identify the differential taste of young households for the above brandy by interacting its product score distance to and the count of other brandies with the share of young consumers in each market. To identify a differential effect of price by income, we construct similar instruments based on the set of products sharing a given product’s price category (cheap vs. expensive) interacted with the share of high-income households in the market.

The mean response across locations to variation in retail prices over time identifies the price coefficient \(\alpha\), exploiting the fact that distillers do not change the wholesale prices \(p^w\) for all products at the same time. We rely on input costs and retail prices in other control states as instruments
to address possible confounding effects of unobserved demand shocks $\xi$ that distillers respond to in setting wholesale price. We identify seasonality and mean preferences $\beta$ for time-invariant product characteristics such as proof and spirit type from systematic variation in market shares of spirits by period or characteristic.

4.5.1 Parameter Estimates

Table 3 presents the demand estimates of our preferred specification of the mixed-logit model. The parameters estimates are precise, and the estimated demand specification captures the patterns of spirit consumption across demographic groups documented in Table 2.

We allow for rich variation across demographics by interacting consumer age and indicators for minority and high educational attainment with proof and indicators for spirit type, bottle size, and import status. The estimates of $\Pi$ reveal significant differences in tastes for spirits across demographic groups. While minority consumers favor brandy, cordials and rum over gin, the reference category, older and college-educated consumers prefer gin to cordials and rum. We also find that older consumers and, to a lesser extent, college-educated consumers are more likely to purchase 1.75 L than 750 ml bottles, our reference category, while minority households are more likely to purchase 375 ml bottles. The estimated demand for wealthier consumers is steeper, which is consistent with the increased consumption of expensive spirits by high-income consumers reported in Table 2. Older and minority consumers favor spirits with higher proof.

We allow for unobserved variation in preferences for a number of the product characteristics, including proof and certain bottle sizes, product categories, and import status. The estimated random coefficients are large, in particular for brandies and for the 375 ml size, indicating that even after controlling for the significant degree of observed differences in tastes on average and by demographic groups, there still exist further similarities between products in these categories that influence their substitution patterns. Lastly, we find that demand increases during the summer and the holiday season and that, on average, consumers prefer products of higher quality and lower proof and favor cordials, rums, and vodkas over gins and brandy.

4.5.2 Implied Upstream Marginal Cost

To consider the response in upstream behavior to alternative retail pricing policies, we require an estimate of the firms’ marginal costs. We combine our demand estimates with the above assumption of Bertrand–Nash pricing to recover the marginal costs that render the observed wholesale prices optimal under the current pricing policy (see the first-order conditions in equation (43)). We rely on these marginal cost estimates in conducting our counterfactual analysis.

We find that the marginal costs of expensive products are on average 2.7 times higher than those of inexpensive products and that brandies and whiskey are the least and most costly products, respectively, to manufacture on average. For the subset of brandies and whiskeys with age information, we find that marginal costs are approximately 1.5 times higher for products that
Table 3: Mixed-Logit Demand

<table>
<thead>
<tr>
<th>Mean Utility</th>
<th>Random Coeff.</th>
<th>Demographic Interactions (II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(β)</td>
<td>(Σ)</td>
<td>AGE</td>
</tr>
<tr>
<td>PRICE</td>
<td>-0.2763</td>
<td></td>
</tr>
<tr>
<td>CONSTANT</td>
<td>-34.8299</td>
<td>0.1759</td>
</tr>
<tr>
<td></td>
<td>(0.8218)</td>
<td>(0.3653)</td>
</tr>
<tr>
<td>375 ml</td>
<td>4.8700</td>
<td>2.1181</td>
</tr>
<tr>
<td></td>
<td>(0.2451)</td>
<td>(0.5896)</td>
</tr>
<tr>
<td>1.75 L</td>
<td>9.0752</td>
<td>0.0204</td>
</tr>
<tr>
<td></td>
<td>(0.2330)</td>
<td>(1.1874)</td>
</tr>
<tr>
<td>BRANDY</td>
<td>-60.4569</td>
<td>10.0606</td>
</tr>
<tr>
<td></td>
<td>(0.3636)</td>
<td>(0.5963)</td>
</tr>
<tr>
<td>CORDIALS</td>
<td>20.7050</td>
<td>0.7215</td>
</tr>
<tr>
<td></td>
<td>(0.3343)</td>
<td>(0.3827)</td>
</tr>
<tr>
<td>RUM</td>
<td>24.8060</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3506)</td>
<td></td>
</tr>
<tr>
<td>VODKA</td>
<td>24.5847</td>
<td>0.0819</td>
</tr>
<tr>
<td></td>
<td>(0.2760)</td>
<td>(0.4193)</td>
</tr>
<tr>
<td>WHISKEY</td>
<td>-0.9444</td>
<td>0.3425</td>
</tr>
<tr>
<td></td>
<td>(0.3187)</td>
<td>(0.3474)</td>
</tr>
<tr>
<td>FLAVORED</td>
<td>-0.6278</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2130)</td>
<td></td>
</tr>
<tr>
<td>IMPORTED</td>
<td>-0.6931</td>
<td>0.4807</td>
</tr>
<tr>
<td></td>
<td>(0.1960)</td>
<td>(0.3356)</td>
</tr>
<tr>
<td>PROOF</td>
<td>-14.6819</td>
<td>0.2244</td>
</tr>
<tr>
<td></td>
<td>(0.7327)</td>
<td>(0.4845)</td>
</tr>
<tr>
<td>QUALITY</td>
<td>4.0281</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.2690)</td>
<td></td>
</tr>
<tr>
<td>HOLIDAY</td>
<td>0.4483</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0075)</td>
<td></td>
</tr>
<tr>
<td>SUMMER</td>
<td>0.0820</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0065)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Robust standard errors are reported in parentheses. Estimates for random coefficients (Σ) and demographic interactions (Π) based on GMM estimation using 2,237,937 observations in 8,470 store-periods and 1,000 simulated agents in each market. AGE is ln(age−20), EDUCATION is an indicator variable equal to one if the agent has some level of college education, INCOME is ln(income), and MINORITY is an indicator equal to one if the agent is non-white. Mean utility contributions of price, holiday, and summer are based on a product fixed effects regression of the product-period fixed effects from the GMM estimation, controlling for price endogeneity. Remaining coefficients result from a projection of the estimated product fixed effects onto time-invariant product characteristics.

distillers age more than four years than for non-aged products. Imported products have 1.8 times the marginal cost of domestically produced products on average, reflecting increased transportation costs and import tariffs that the distillers pay.

Table 4 documents the significant market power implied by our cost estimates; on average, the firms earn 36.5 cents in profit per dollar in revenue. Products manufactured by larger firms, such as Diageo and Bacardi, have lower Lerner indices (roughly 34%) while smaller manufacturers such as Jacquin and Sazerac operate niche product portfolios on which they keep 46.5% and 42.3% of revenue in profit, respectively. Across manufacturers, CHEAP and 375 ml products are particularly profitable.
The presence of upstream market power raises two points. First, current policy may implicitly redistribute rents between firms, or between firms as a group and consumers or the PLCB. Second, upstream firms possess the ability to respond to changes in policy – a factor that we must account for in considering the implications of any alternative policy.

### 4.6 Third-Degree Price Discrimination vs. Uniform Pricing

In this section we use the estimates of consumer demand and upstream marginal cost to evaluate the quantitative implications of 3DPD taking into account the upstream firm response of distillers with market power. In Section 2 we demonstrated that the downstream firm set 3DPD retail prices based on the product-level demand curvature (σ) of products in the portfolio. Thus, 3DPD profits depend upon heterogeneity of these demand curvatures. In Figure 5 we document that our demand estimates do indeed generate differentiation in demand curvature across the state for all products (panel a) as well as for a single product (panel b). Also note that our estimated demand system generates an average demand curvature of $\hat{\sigma} = 1.035$ and 96.1% of products have demand curvature greater than one. Thus, for the majority of products we consider, consumer demand is not log-concave. Quint (2014, §4.3) shows that many demand systems (e.g., multinomial and nested logit) impose log-concavity by assumption whereas the random coefficient logit demand model does not. From Figure 5 we also observe that the distribution of demand curvature does not follow any obvious geographic patterns. This suggests that pricing strategies which treat, say, Philadelphia and Pittsburgh differently than the rest of the state are not likely to generate much additional downstream profit.

Recent focus has been placed on the use of sufficient statistics to infer the equilibrium implications of counterfactual policies (e.g., Saez, 2001; Chetty, 2009a; Chetty, 2009b). In industrial

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>DIA GEO</th>
<th>BACARDI</th>
<th>BEAM</th>
<th>JACQUIN</th>
<th>SAZERAC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>By Spirit Type:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRANDY</td>
<td>45.25</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>50.73</td>
<td>-</td>
</tr>
<tr>
<td>CORDIALS</td>
<td>35.79</td>
<td>38.87</td>
<td>17.19</td>
<td>47.63</td>
<td>57.55</td>
<td>38.69</td>
</tr>
<tr>
<td>GIN</td>
<td>37.07</td>
<td>32.52</td>
<td>20.37</td>
<td>39.87</td>
<td>37.80</td>
<td>54.03</td>
</tr>
<tr>
<td>RUM</td>
<td>37.36</td>
<td>32.25</td>
<td>38.06</td>
<td>40.81</td>
<td>48.30</td>
<td>-</td>
</tr>
<tr>
<td>VODKA</td>
<td>35.82</td>
<td>38.58</td>
<td>-</td>
<td>37.33</td>
<td>38.81</td>
<td>43.01</td>
</tr>
<tr>
<td>WHISKEY</td>
<td>34.30</td>
<td>32.01</td>
<td>18.86</td>
<td>36.55</td>
<td>36.17</td>
<td>39.47</td>
</tr>
<tr>
<td><strong>By Price:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHEAP</td>
<td>45.70</td>
<td>46.45</td>
<td>44.46</td>
<td>45.53</td>
<td>46.48</td>
<td>44.26</td>
</tr>
<tr>
<td>EXPENSIVE</td>
<td>27.12</td>
<td>28.92</td>
<td>24.48</td>
<td>27.57</td>
<td>-</td>
<td>26.52</td>
</tr>
<tr>
<td><strong>By Bottle Size:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>375 ml</td>
<td>59.18</td>
<td>59.57</td>
<td>65.96</td>
<td>72.07</td>
<td>91.76</td>
<td>47.71</td>
</tr>
<tr>
<td>750 ml</td>
<td>36.99</td>
<td>34.16</td>
<td>33.78</td>
<td>44.45</td>
<td>56.80</td>
<td>54.25</td>
</tr>
<tr>
<td>1.75 L</td>
<td>29.16</td>
<td>23.19</td>
<td>21.43</td>
<td>29.07</td>
<td>35.21</td>
<td>37.12</td>
</tr>
<tr>
<td><strong>ALL PRODUCTS</strong></td>
<td>36.49</td>
<td>34.25</td>
<td>34.99</td>
<td>39.43</td>
<td>46.48</td>
<td>42.33</td>
</tr>
</tbody>
</table>

Notes: Table displays average Lerner index, defined as $100 \times (p^w - \hat{c})/p^w$, weighted by bottles sold.
organization our theories indicate that estimated demand elasticities are often useful towards informing the implications of mergers and government policy. In Section 2, however, we demonstrate that knowledge of demand elasticities for most demand systems provides little insight into the implications of 3DPD. As upstream firms set optimal prices using such elasticities, however, the degree to which downstream demand curvature and upstream product-level demand elasticities coincide may be useful towards understanding whether the strategic response of upstream firms is quantitatively important to the downstream profitability of 3DPD. In Figure 6 we document that there exists no connection between estimated demand curvature and elasticity. This indicates that the latter provides no information towards understanding the implications of 3DPD and that the upstream firm response is not likely to have a large effect on the profitability of downstream 3DPD.

**Figure 5: Evidence for Third-Degree Price Discrimination**

Notes: Maps correspond to the spatial distribution of demand curvature (i.e., $\sigma$ in the text) in Pennsylvania during the sample. Outlined polygons correspond to geographic markets (i.e., “stores” in the text). Markets are broken up into terciles according to their average (weighted by quantity sold) demand curvature. In panel (a) we assigned markets into each bin according the average demand curvature of all products. In panel (b) we limited the analysis to 750 ml bottle of Jack Daniels Old No. 7 Black Label, a popular whiskey product. Sales-weighted average values for each bin are indicated in parentheses. All statistics based on 2003.

**Figure 6: Demand Elasticities as Sufficient Statistic?**

Notes: Maps correspond to the spatial distribution of demand curvature $\hat{\sigma}$ in Pennsylvania during the sample. Outlined polygons correspond to geographic markets (i.e., “stores” in the text). In panel (a) we compare estimated demand curvature $\hat{\sigma}$ and downstream demand elasticity $\hat{\varepsilon}(p)$ for all products. In panel (b) we limited the analysis to 750 ml bottle of Jack Daniels Old No. 7 Black Label, a popular whiskey product. All statistics based on estimated for 2003.
4.6.1 Generating the Right Benchmark

Recall that the consumer demand and upstream marginal cost estimates of Section 4.5.1 are based on an equilibrium in which a government agency, the PLCB, chooses a uniform tax rate (30%) and upstream firms respond by choosing product-level, statewide wholesale prices. In constructing our baseline, we use the estimated consumer demand and marginal cost estimates to construct a baseline Stackelberg equilibrium where upstream firms first choose product-level, statewide wholesale prices taking into account the statewide retail prices set by the downstream monopolist. We call this Stackelberg equilibrium “Uniform.” To address the equilibrium implications of 3DPD, we compare this equilibrium to two alternative equilibria. In the first, we allow the downstream monopolist to set product-level prices (312) by store (456) in each period but do hold the upstream wholesale prices fixed at the levels from the “Uniform” Stackelberg equilibrium. We call this equilibrium “Upstream Fixed” and we note that comparison between this equilibrium and the “Uniform” equilibrium is the multi-product analog to the analysis of Della Vigna and Gentzkow (2017). In our second equilibrium we solve for the Stackelberg equilibrium where the upstream firms set prices in a Bertrand-Nash pricing equilibrium taking into account the downstream pricing response of the monopolist.15 We call this Stackelberg equilibrium “Upstream Flexible.” To ease the computational burden of computing roughly 150,000 3DPD retail prices each period, we consider only four pricing periods in 2003 where each period corresponds to one quarter of the year (i.e., periods included correspond to January, April, September, and December).

Throughout our analysis, we assume that the cost of implementing 3DPD (F) is sufficiently small such that the monopolist always chooses this pricing strategy. We then evaluate the degree to which the upstream firm response influences our estimate of F in the event we do not observe 3DPD in the data. That is, we evaluate whether ignoring strategic interactions along the vertical market structure materially impacts our estimates of the minimum F required to make 3DPD not profitable.

4.6.2 Results

In Table 5, we compare the equilibrium implications of ignoring the strategic behavior of the upstream firms.

We find that 3DPD increases downstream profits a modest 2.43% at the set of equilibrium wholesale prices from the “Uniform” pricing Stackelberg equilibrium. Allowing for the upstream firms to modify their wholesale prices and resolving the Stackelberg equilibrium leads to a reduction in downstream profits from $43.84 to $43.79 million. Consumers are generally worse off under 3DPD, though the upstream response to increase 66.9% of wholesale prices leads to an increase in the retail price of nearly half of the products and a general increase in average retail price. Thus,

15We do so by totally-differentiating the downstream firm first-order conditions when the downstream firm sets product-level prices which vary by store but face wholesale costs which only vary by product. See Villas-Boas (2007) for the case when the downstream firm does not 3DPD.
Table 5: Equilibrium Effects of Price Discrimination with Upstream Market Power (Random Coefficient Logit Demand)

<table>
<thead>
<tr>
<th></th>
<th>Uniform</th>
<th>Upstream Fixed</th>
<th>Upstream Flexible</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demand Curves</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outside Elasticity</td>
<td>-3.06</td>
<td>-3.07</td>
<td>-3.07</td>
</tr>
<tr>
<td>Demand Elasticity ($\hat{\varepsilon}$)</td>
<td>-3.96</td>
<td>-3.97</td>
<td>-3.97</td>
</tr>
<tr>
<td>Curvature ($\hat{\sigma}$)</td>
<td>1.03</td>
<td>1.03</td>
<td>1.03</td>
</tr>
<tr>
<td><strong>Wholesale Prices ($p^w$)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>10.34</td>
<td>10.34</td>
<td>10.35</td>
</tr>
<tr>
<td>% Increase</td>
<td>N/A</td>
<td>0.00</td>
<td>66.90</td>
</tr>
<tr>
<td><strong>Retail Prices ($p^r$)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>16.20</td>
<td>16.33</td>
<td>16.34</td>
</tr>
<tr>
<td>% Increase</td>
<td>N/A</td>
<td>46.96</td>
<td>47.09</td>
</tr>
<tr>
<td><strong>Profits</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wholesale ($\Pi^w$)</td>
<td>29.34</td>
<td>29.86</td>
<td>29.89</td>
</tr>
<tr>
<td>Retail ($\Pi^r$)</td>
<td>42.79</td>
<td>43.84</td>
<td>43.79</td>
</tr>
<tr>
<td><strong>Quantity Sold ($q$)</strong></td>
<td>7.18</td>
<td>7.31</td>
<td>7.31</td>
</tr>
<tr>
<td><strong>Consumer Welfare</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compensating Variation ($CV$)</td>
<td>N/A</td>
<td>1.09</td>
<td>1.13</td>
</tr>
<tr>
<td>% Residents Better-Off ($CV &lt; 0$)</td>
<td>N/A</td>
<td>65.96</td>
<td>65.81</td>
</tr>
</tbody>
</table>

Notes: Table presents descriptive statistics across equilibria described in the text using demand estimates presented in Table 3 and estimates of upstream marginal cost presented in Miravete et al. (2018b). “Outside Elasticity,” “Demand Elasticity,” and “Curvature” are the elasticity of spirits as a category, the average product-level (i.e., jlt) downstream demand elasticities, and the average downstream product-level (i.e., jlt) demand curvature, respectively. “% Increase” is the share of wholesale or retail prices which increase relative to “Uniform” pricing. Compensating variation is in millions and is calculated following Section 4.3.1 where product-level pricing (i.e., “Uniform”) is the benchmark equilibrium. “% Residents Better-Off” is the share of PA residents which have compensating variation less than zero (i.e., they would be willing to pay to enter into the new equilibrium) from 3DPD. For reference, tax revenue in the data when the PLCB acts as a government agency and sets a uniform 30% tax rate is 47.68, or 11.4% greater than under the “Uniform” pricing Stackelberg equilibrium.

The wholesale price response makes consumers worse off, reflected by an increase in compensating variation; i.e., an increase in the amount of money one would have to give them to make them indifferent between the “Upstream Flexible” and “Uniform” equilibria. Not all consumers, however, are worse off as 65.81% of Pennsylvania residents prefer 3DPD even after accounting for upstream pricing.

In Figure 7 we evaluate the implications of upstream firm pricing on downstream firm profits and our estimate of the costs of implementing 3DPD. Define downstream “capture” as the share of potential profits attained by 3DPD, i.e.,

$$\frac{\Pi^{‘Flexible’} - \Pi^{‘Uniform’}}{\Pi^{‘Fixed’} - \Pi^{‘Uniform’}}$$
where $\Pi^{\text{Uniform}}$, $\Pi^{\text{Fixed}}$, and $\Pi^{\text{Flexible}}$ refer to total period $t$ downstream profits (summed across all local stores) in the “Uniform,” “Upstream Fixed,” and “Upstream Flexible” equilibria, respectively. For values of this statistic close to one, the pricing behavior of upstream firms has little impact on downstream profits whereas for values close to zero the behavior of upstream firms has a significant impact on the downstream profitability of 3DPD.

**Figure 7: Upstream Undoing**

![Graph showing realized change in downstream profit and estimate of critical value](image)

Notes: Figures describe the implications of ignoring the pricing behavior of upstream firms. Panel a documents that upstream firms limit the gains of 3DPD. Panel b documents that ignoring upstream pricing leads the researcher to over-estimate the costs required to generate uniform pricing.

In panel a, we document that the downstream firm is able to achieve the lion’s share of 3DPD, between 92.5% and 96.9% of the potential profits. Given that the downstream firm will choose to not 3DPD for any $F > 0$ such that

$$\Pi^{3DPD} - F \leq \Pi^{\text{Uniform}},$$

we can define $F^*$ as the value of $F$ such that (45) hold with equality. When upstream pricing decreases downstream profits (i.e., $\Pi^{\text{Flexible}} < \Pi^{\text{Fixed}}$), ignoring the behavior of upstream firms will lead to an over-estimated of $F^*$. The question then is whether using $\Pi^{\text{Fixed}}$ or $\Pi^{\text{Flexible}}$ materially impacts our estimate of $F^*$. In panel (b), we document that this over-estimate ranges between 3.2 and 8.1%.

Our analysis thus far evaluates only an extreme form of 3DPD where the downstream firm is able to choose local prices for 456 stores. In reality, consumer arbitrage would likely erode profits under such a policy, particularly in markets where stores are close to each other such as the urban areas of Philadelphia and Pittsburgh. Interestingly, we find that the results in Figure 7 hold-up when we consider less coarse policies such as where the downstream firm chooses retail prices in Philadelphia, Pittsburgh, and the rest of the state (3 local markets versus 456 local markets) though the benefits of 3DPD are much smaller so the magnitude of $\Pi^{3DPD}$ and $F^*$ are less than in the case of 456 local markets.
Figure 8: Upstream Undoing When Retailer Uses a Simpler 3DPD Strategy

Notes: Figures describe the implications of ignoring the pricing behavior of upstream firms when the downstream firm chooses prices for Pittsburgh, Philadelphia, and the rest of the state (3 zones). Panel a documents that upstream firms limit the gains of 3DPD. Panel b documents that ignoring upstream pricing leads the researcher to over-estimate the costs required to generate uniform pricing.

5 Concluding Remarks

To be continued...
References


Appendix

A Uniform Retail Pricing

We address how the equilibrium varies with the degree of curvature of demand, $\sigma$. Figure A.1 (a) shows that total output per market decreases at a nearly linear rate with $\sigma$, as given by (21a), to the limiting value of $q_u(1) = e^{-2}$. Figure A.1 (b) depicts the optimal retail and wholesale prices, (21b) and (20), respectively, when retail pricing is uniform across local markets with identical demand curvature $\sigma$. Both prices decrease with $\sigma$ with $p_u(\sigma) = c_u(\sigma)(3 - \sigma)/(2 - \sigma)$ ranging from being 50\% larger to double $c_u(\sigma)$ when $\sigma = 0$ and $\sigma = 1$, respectively.

Figure A.1: Uniform Pricing: Sales and Prices

![Figure A.1](image)

(a) Sales per Market  (b) Wholesale and Retail Prices

Figure A.2: Uniform Pricing: Retail and Wholesale Profits

![Figure A.2](image)

Figure A.2 shows the wholesale and retail profits when double marginalization is fully anticipated, i.e., equations (22) and (23b), respectively, where the former exceeds the later by a factor of $(2 - \sigma)$. 
B Price Discrimination with Bulow-Pfleiderer’s Constant Curvature Demand

Demand function (17) is log-concave, it is always decreasing, linear when \( \sigma = 0 \), and convex for \( 0 \leq \sigma \leq 1 \), leading to an incomplete constant pass-through ratio \( \rho = 1/(2 - \sigma) \).

For the the numerical analysis that follows we normalize wholesale cost \( \zeta = 0 \). We also assume \( a^w = a^s = 1 \) so that the maximum willingness to pay in each market are the same. Making \( b^w = b^s = 1 \) we normalize the size \( (a^i/b^i)^{1/(1-\sigma^i)} \) of both market types. All these normalizations reduce the key difference among markets to the convexity of their inverse demands so that \( \sigma^w \geq \sigma^s \geq 0 \). Thus, we write \( \gamma = \sigma^s/\sigma^w \in [0, 1] \).

Figure 3 explores the effect of curvature dissimilarity across different compositions of weak and strong markets in the equilibrium outcome of 3DPD after solving for \( c^* \) in the following normalized version of (28):

\[
c^*(\sigma^w, \gamma) = \frac{\frac{1}{2 - \sigma^w}}{\frac{1}{1 - \sigma^w} + \frac{1}{1 - \gamma \sigma^w} \gamma \sigma^w}. \tag{B.1}
\]

C Retail Price Discrimination

Notice in equation (24) that for any wholesale price \( c \), the retail profit maximization solution at each local market is given by expressions analogous to (18a)-(18b) after replacing \( \{a, b, \sigma\} \) by \( \{a^w, b^w, \sigma^w\} \) or \( \{a^s, b^s, \sigma^s\} \) for weak and strong markets, respectively.

The equilibrium wholesale price with retail 3DPD, \( c^*(\sigma^w, \sigma^s) \) falls between \( c^u(\sigma^w) \) and \( c^u(\sigma^s) \) because \( \min\{-q^u(c^*, \sigma^w)/q^u(c, \sigma^w)\} \leq c^*(\sigma^w, \sigma^s) - \zeta \leq \max\{-q^u(c^*, \sigma^s)/q^u(c^*, \sigma^s)\} \) and:

\[
- \frac{q^i(c^*, \sigma^i)}{q^u(c^*, \sigma^i)} = \left[ \frac{a^i - c^*}{b^i(2 - \sigma^i)} \right] \frac{1}{1 - \sigma^i} = (1 - \sigma^i)(a^i - c^*), \quad \text{for } i = w, s. \tag{C.1}
\]

Since condition \( \sigma^w \geq \sigma^s \geq 0 \) is needed for retail price discrimination to increase overall sales, it must be the case that \( (1 - \sigma^w)(a^w - c^*) \leq (1 - \sigma^s)(a^s - c^*) \) when \( a^w \leq a^s \). Then,

\[
c^u(\sigma^w) = \frac{(1 - \sigma^w)a^w + \zeta}{2 - \sigma^w} \leq c^*(\sigma^w, \sigma^s) \leq \frac{(1 - \sigma^s)a^s + \zeta}{2 - \sigma^s} = c^u(\sigma^s). \tag{C.2}
\]
Table C.1: Uniform Pricing vs. 3DPD Performance Measures

<table>
<thead>
<tr>
<th></th>
<th>γ = 0.1</th>
<th>γ = 0.5</th>
<th>γ = 0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Naïve Sales Ratio</td>
<td>1.4188</td>
<td>1.3017</td>
<td>1.1112</td>
</tr>
<tr>
<td></td>
<td>[1.0428 , 1.9849]</td>
<td>[1.0242 , 1.7721]</td>
<td>[1.0049 , 1.3130]</td>
</tr>
<tr>
<td>(b) Actual Sales Ratio</td>
<td>1.0158</td>
<td>0.9947</td>
<td>0.9928</td>
</tr>
<tr>
<td></td>
<td>[1.0176 , 0.8227]</td>
<td>[1.0100 , 0.7753]</td>
<td>[1.0020 , 0.9904]</td>
</tr>
<tr>
<td>(c) Wholesale Price, c*</td>
<td>0.4414</td>
<td>0.3879</td>
<td>0.3230</td>
</tr>
<tr>
<td></td>
<td>[0.4859 , 0.4708]</td>
<td>[0.4807 , 0.2129]</td>
<td>[0.4753 , 0.1280]</td>
</tr>
<tr>
<td>(d) Price Difference Ratio</td>
<td>0.8325</td>
<td>0.8999</td>
<td>0.9799</td>
</tr>
<tr>
<td></td>
<td>[0.9772 , 0.5827]</td>
<td>[0.9872 , 0.7463]</td>
<td>[0.9974 , 0.9600]</td>
</tr>
<tr>
<td>(e) 3DPD to Uniform Profit (Naïve)</td>
<td>6.3686</td>
<td>4.3565</td>
<td>1.7514</td>
</tr>
<tr>
<td></td>
<td>[1.0705 , 8.2967]</td>
<td>[1.0394 , 5.4766]</td>
<td>[1.0079 , 1.9320]</td>
</tr>
<tr>
<td>(e) 3DPD to Uniform Profit (Actual)</td>
<td>2.9991</td>
<td>2.3770</td>
<td>1.3625</td>
</tr>
<tr>
<td></td>
<td>[1.0294 , 2.5033]</td>
<td>[1.0116 , 2.1881]</td>
<td>[1.0024 , 1.4273]</td>
</tr>
<tr>
<td>(f) Profit Ratio</td>
<td>0.6838</td>
<td>0.7602</td>
<td>0.9033</td>
</tr>
<tr>
<td></td>
<td>[0.9537 , 0.3025]</td>
<td>[0.9735 , 0.4020]</td>
<td>[0.9945 , 0.7404]</td>
</tr>
</tbody>
</table>

Notes: Mean value of each performance item; 90% confidence intervals located in brackets.