

NBER WORKING PAPER SERIES

LEARNING FIRM CONDUCT:  
PASS-THROUGH AS A FOUNDATION FOR INSTRUMENT RELEVANCE

Adam Dearing  
Lorenzo Magnolfi  
Daniel Quint  
Christopher J. Sullivan  
Sarah B. Waldfogel

Working Paper 32863  
<http://www.nber.org/papers/w32863>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
August 2024

Some results first appeared in the working paper “Falsifying Firm Conduct,” which this paper supersedes. We thank Anya Tarascina for tremendous research assistance, and Brett Hollenbeck, Nathan Seegert, and Kosuke Uetake for kindly sharing data with us. We also thank Christopher Conlon, Mert Demirer, Paul Grieco, Philip Haile, Justin Johnson, Nathan Miller, Conor Ryan, Katja Seim, and Matthew Weinberg for helpful discussions, as well as numerous seminar and conference participants. Any errors are our own. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2024 by Adam Dearing, Lorenzo Magnolfi, Daniel Quint, Christopher J. Sullivan, and Sarah B. Waldfogel. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Learning Firm Conduct: Pass-Through as a Foundation for Instrument Relevance

Adam Dearing, Lorenzo Magnolfi, Daniel Quint, Christopher J. Sullivan, and Sarah B. Waldfogel  
NBER Working Paper No. 32863

August 2024

JEL No. L0

**ABSTRACT**

Researchers often test firm conduct models using pass-through regressions or instrumental variables (IV) methods. The former has limited applicability; the latter relies on potentially irrelevant instruments. We show the falsifiable restriction underlying the IV method generalizes the pass-through regression, and cost pass-through differences are the economic determinants of instrument relevance. We analyze standard instruments' relevance and link instrument selection to target counterfactuals. We illustrate our findings via simulations and an application to the Washington marijuana market. Testing conduct using targeted instruments, we find the optimal ad valorem tax closely matches the actual rate.

Adam Dearing  
SC Johnson Graduate School of Management  
SC Johnson College of Business  
Cornell University  
369 Sage Hall  
114 E. Ave.  
Ithaca, NY 14853  
and NBER  
aed237@cornell.edu

Lorenzo Magnolfi  
University of Wisconsin-Madison  
1180 Observatory Drive  
Madison, WI 53706-1393  
magnolfi@wisc.edu

Daniel Quint  
University of Wisconsin-Madison  
1180 Observatory Drive  
Madison, WI 53706-1393  
dquint@ssc.wisc.edu

Christopher J. Sullivan  
University of Wisconsin-Madison  
1180 Observatory Drive  
Madison, WI 53706-1393  
cjsullivan@wisc.edu

Sarah B. Waldfogel  
Darla Moore School of Business  
University of South Carolina  
waldfogs@mailbox.sc.edu

# 1 Introduction

Understanding how firms behave is a central objective in economics research. As firm conduct is frequently unknown, researchers have developed various methods to infer it from data. An influential idea in the early empirical industrial organization (IO) literature was to utilize the implications of cost pass-through rates to test firm conduct.<sup>1</sup> The simplest implementation in [Sumner \(1981\)](#) estimates pass-through from a regression of prices on unit taxes. Because different oligopoly models predict different pass-through levels (e.g., [Weyl and Fabinger, 2013](#)), the estimates from these regressions reject incorrect models of firm behavior. This approach has been applied across multiple fields of economics, including IO, trade, and public finance.<sup>2</sup>

As noted in the literature (e.g., [Bulow and Pfleiderer, 1983](#); [MacKay, Miller, Remer, and Sheu, 2014](#)), the pass-through regression approach to infer conduct is only valid in a narrow class of models and when a special cost shifter (e.g., unit tax or wholesale price) is observed in the data. An alternative approach, originating with [Bresnahan \(1982\)](#) and [Lau \(1982\)](#), relies on exogenous variation in market conditions to distinguish models. In a differentiated products environment, [Berry and Haile \(2014\)](#) formalize this intuition into a falsifiable restriction implemented with a set of potentially exogeneous instruments. While their approach is widely applicable, it is crucial to assess the relevance of instruments for falsification, i.e., their ability to distinguish misspecified models from the truth.

In this paper, we develop a framework for understanding the economic determinants of instrument relevance when falsifying models of firm conduct. We find that differences in particular features of cost pass-through matrices across oligopoly models underpin falsifiability. This has several practical implications for learning conduct, of which we highlight three.

First, we show that the falsifiable restriction in [Berry and Haile \(2014\)](#) coincides with the regression approach in [Sumner \(1981\)](#) in the limited class of models with constant pass-through – the only class where the regression approach is valid ([MacKay et al., 2014](#)). Thus, the instrument-based approach generalizes the pass-through regression approach.

Second, we use our framework to evaluate the relevance of various conduct instruments. In general, we find that rival cost shifters, product characteristics, and tax rates have different empirical content; each targets different features of the model’s pass-through matrix and other characteristics. We characterize exactly the aspects of conduct each instrument targets.

Third, we show how this characterization can guide instrument selection for particular counterfactual analyses. For example, computing a counterfactual Laffer curve to determine

---

<sup>1</sup>See, for example, [Sumner \(1981\)](#); [Sullivan \(1985\)](#); [Panzar and Rosse \(1987\)](#).

<sup>2</sup>For recent applications of this method see, e.g., [Pless and Van Benthem \(2019\)](#) and [Brugués and De Simone \(2024\)](#).

a revenue-maximizing tax rate depends on both demand and firm conduct (e.g., [Miravete, Seim, and Thurk, 2018](#); [Hollenbeck and Uetake, 2021](#)). The features of conduct that determine the optimal tax rate are exactly those targeted by an instrument based on past variation in that tax rate. When such variation is not available, our framework illustrates the trade-offs between other possible instruments.

While our falsification results are presented in a population context, they directly inform finite sample testing procedures based on the [Rivers and Vuong \(2002\)](#) (RV) model selection test. Lack of falsification due to irrelevant instruments leads to degeneracy of the RV test and unreliable inference ([Duarte, Magnolfi, Sølvesten, and Sullivan, 2024](#)). Through Monte Carlo simulations, we illustrate the usefulness of our framework for testing models of conduct with RV, highlighting the importance of ex-ante instrument selection for both testing specific models and informing counterfactual outcomes. Our results also reveal the economic determinants of the ex-post weak instrument diagnostic in [Duarte et al. \(2024\)](#).

As an application of our results, we consider the problem of setting the optimal ad valorem tax rate in the Washington State marijuana market in 2016. Tax revenue is a key consideration for governments legalizing marijuana, and [Hollenbeck and Uetake \(2021\)](#) show the importance of accounting for firm market power to determine the optimal tax in this market. Following legalization, it is unclear to what extent retailers behave strategically. [Escudero \(2018\)](#) suggests that immediately following legalization (2014-2016), retailers may have used a simple rule of setting prices equal to twice marginal costs, a strategy referred to as Keystone pricing after a jewelers’ trade publication in which it was widely promoted in the late 1800s. [Hollenbeck, Hristakeva, and Uetake \(2024\)](#) provide reduced-form evidence of a more nuanced relationship between markups and the competitive environment in later years (2018-2022).

Motivated by these findings, we use our framework to formally test the strategically sophisticated Bertrand-Nash pricing model against the simpler Keystone pricing model in a data sample from 2014-2017. These two models reflect two “extremes” of sophistication in how firms respond to competition, and it is *a priori* unclear which will better fit the data. We first estimate a demand system similar to [Hollenbeck and Uetake \(2021\)](#). We then test the Bertrand-Nash model against Keystone pricing using ad valorem tax instruments, which our framework shows are relevant for testing these models and will falsify models that yield incorrect optimal tax predictions. Our test indicates that Keystone provides a better fit than Bertrand-Nash during our sample period. Based on these findings, we solve for the optimal state-level ad valorem tax in 2016. The revenue maximizing state-level tax rate under Keystone is 38%, almost identical to the 37% rate the government adopted in 2015. Alternative conduct models imply different optimal taxes.

This paper complements existing work ([Backus, Conlon, and Sinkinson, 2021](#); [Duarte](#)

et al., 2024) that advocates a model selection approach and offers inferential advances in conduct testing. The falsification framework we develop can help clarify the economic determinants of instrument strength for the RV test and inform ex ante instrument selection. We illustrate this connection in our Monte Carlo simulations and our empirical application, which both implement testing with RV.

Our work is also related to recent work on demand curvature and pass-through. Weyl and Fabinger (2013) link pass-through to both firm conduct and demand curvature. Miravete, Seim, and Thurk (2024) show how standard discrete choice demand models can be made flexible, so as to produce a full range of demand curvature. Our results underscore the need for flexible demand specifications that can accurately capture pass-through patterns.

We also contribute to a long line of literature on optimal tax policy, in which the Laffer curve and other aspects of tax incidence have been studied extensively. Following the seminal contribution of Ramsey (1927), this literature has traditionally imposed the assumption of perfect competition (Auerbach, 1985, 2002; Slemrod, 1990). However, there is extensive theoretical evidence that conduct has important implications for tax policy. For example, Suits and Musgrave (1953) and Anderson, de Palma, and Kreider (2001) show that the relative efficiency of unit and ad valorem taxation depends on the form of market conduct. Recent work by Kroft, Laliberté, Leal-Vizcaíno, and Notowidigdo (2023) and Brugués and De Simone (2024) attempts to perform inference on conduct to better-inform tax policy. These papers use “conduct parameter” and “conjectural variations” approaches; like pass-through regressions, these approaches are only micro-founded for a narrow range of models (Makowski, 1987; Lindh, 1992; Corts, 1999). We contribute to this literature by providing guidance on which features of conduct – captured by the pass-through matrix – are most relevant to tax policy. Our insights can help guide instrument choice when prior variation in tax rates is unavailable. Relative to approaches based on pass-through regressions and conjectural variations, our approach is valid in a much broader range of settings, including those used in state-of-the-art structural equilibrium models.

The paper proceeds as follows. Section 2 describes the environment and shows that the exclusion restriction approach generalizes the pass-through approach. Section 3 presents our main results showing how differences in pass-through matrices form the economic determinants of falsifiability. Section 4 compares the relevance of demand side, cost side, and tax instruments. Section 5 motivates instrument selection for optimal tax counterfactuals. In Section 6, we perform Monte Carlo simulations to show the applicability of our results to a setting using the RV test. Section 7 develops an application to the Washington marijuana market. Section 8 concludes. Proofs are found in Appendix A.

## 2 Two Approaches to Falsification

We consider falsification of models of firm conduct using data across many markets.<sup>3</sup> A set of multi-product firms compete in each market  $t$ ; for simplicity, we assume the same set of  $J$  products is sold in every market, although their characteristics may differ across markets. For each product and market combination  $(j, t)$ , the researcher observes price  $p_{jt}$ , market share  $s_{jt}$ , a vector of product characteristics  $x_{jt}$ , and a vector of cost shifters  $w_{jt}$  that affects the product's marginal cost. For any variable  $a_{jt}$ , let  $a_t$  denote the vector of values for all products  $j$  in market  $t$ . We assume that, for all markets  $t$ , the demand system is  $s_t = \mathcal{d}(p_t, x_t, \xi_t)$ , where  $\xi_t$  is a vector of unobserved product characteristics. To focus on the supply side, we assume that the demand system is already known to the researcher. We normalize market size to 1, so that quantity  $q_{jt}$  and market share  $s_{jt}$  can be used interchangeably.

The data in each market  $t$  are generated by equilibrium play in some true model of firm behavior, characterized by a system of first-order conditions,

$$p_t = \Delta_{0t} + c_{0t}, \quad (1)$$

where  $\Delta_{0t}$  is the true vector of markups in market  $t$  and  $c_{0t}$  is the true vector of marginal costs. Because true markups (or, equivalently, true costs) are unobserved, we consider theoretically plausible models to be falsified by the data. Under any imposed model  $m$ , we can calculate the implied markups  $\Delta_{mt}$  as a known function of observables and demand primitives. Implied marginal costs  $c_{mt}$  can therefore also be calculated via a model-specific version of the first-order conditions in Equation (1), as  $p_t - \Delta_{mt}$ .

For the first-order conditions of any model  $m$  to characterize a well-defined empirical model, we require the following, analogous to Assumption 13 in [Berry and Haile \(2014\)](#):

**Assumption 1.** (*Equilibrium Uniqueness*) For any model  $m$ , including the true model, either there exists a unique equilibrium, or the equilibrium selection rule is such that the same  $p_t$  arises whenever the vector  $(c_{mt}, x_t, \xi_t)$  is the same.

Several methods have been proposed to learn whether the imposed model  $m$  is falsified by the data. We focus on two in the next subsections: the pass-through approach in [Sumner \(1981\)](#), and the exclusion restriction approach in [Berry and Haile \(2014\)](#).

---

<sup>3</sup>Falsification encompasses cases of nested and non-nested models. We focus on the latter. For the former, researchers can pursue identification of conduct; we discuss the usefulness of our framework for this approach in [Appendix B](#).

## 2.1 Pass-through Regressions to Learn Conduct

Early tests of market power (e.g., [Sumner, 1981](#)) relied on a specific implication of different models of conduct: cost (or, equivalently, unit-tax) pass-through. In a simple class of “rule-of-thumb” pricing models, where firms set prices as a multiple of cost, own cost pass-through can be expressed as a constant scalar  $\phi_m$  across products and markets. Marginal cost pricing ( $\phi_m = 1$ ) and Keystone pricing ( $\phi_m = 2$ ) are examples in this class. If data are generated by a model  $m = 0$  in this class, Equation (1) becomes  $p_{jt} = \phi_0 c_{0jt}$ . Because cost pass-through is represented by a scalar parameter, it can be measured via linear regression. To implement the regression, we assume that marginal costs are constant in quantity, linear in observable cost shifters  $w_{jt}$ , and additively separable in an unobserved shock, or  $c_{0jt} = w_{jt}\gamma + \omega_{0jt}$ . While the researcher can control for  $w_{jt}$ , she cannot keep  $\omega_{0jt}$  fixed across markets.

With these assumptions, firms set prices  $p_{jt} = \phi_0 (w_{jt}\gamma + \omega_{0jt})$ . If we knew the value of  $\gamma$ , then the true own pass-through  $\phi_0$  could be consistently estimated by regressing  $p_{jt}$  on  $w_{jt}$ . While  $\gamma$  is unknown for a vector of generic cost shifters, there may be components of cost for which  $\gamma = 1$ . These include unit taxes which are levied on the firm ( $\tau_{jt}$ ) and wholesale prices ( $p_{jt}^W$ ). As in [Sumner \(1981\)](#), suppose we observe a market-level unit tax  $\tau_t$  for which there is variation across markets. We can residualize all variables, including  $\tau_t$  with respect to the observed cost shifters  $w_{jt}$ , excluding  $\tau_t$ . Then, we can estimate own cost pass-through by regressing  $p_{jt}$  on  $\tau_t$ :  $\phi_0 = \beta_{PTR} = \text{cov}(\tau_t, p_{jt})/\text{var}(\tau_t)$ . Hence, when  $\beta_{PTR} \neq \phi_m$ , one can falsify the rule-of-thumb model for which the firm prices at  $\phi_m$  times marginal cost. For example, if  $\beta_{PTR} \neq 1$ , marginal cost pricing is falsified.

However, the pass-through regression approach has a few conceptual and practical issues. First, the required variation in taxes or observed cost components may not be available in the data. Second, outside of constant pass-through environments, such as the one described above or the homogenous product environments in [Bulow and Pfleiderer \(1983\)](#), own and rival cost pass-through are not structural parameters. Instead, in market  $t$ , cost pass-through under model  $m$  is represented by the  $J \times J$  matrix  $P_{mt}$ , which varies with realizations of both observable and unobservable determinants of demand and cost. Consequently, it is generally not possible to recover the elements of  $P_{mt}$  via regression ([MacKay et al., 2014](#)).

## 2.2 Exclusion Restrictions to Learn Conduct

Following the seminal work of [Bresnahan \(1982\)](#) and [Lau \(1982\)](#), an alternative approach to learn conduct in the presence of unobservable shocks to demand and cost is to leverage exogenous variation in market conditions. [Berry and Haile \(2014\)](#) generalized this intuition: if a researcher can construct instruments  $z_{jt}$  that are mean independent of the unobserved cost

shocks under the true model, the exclusion restriction can serve as a falsifiable restriction.

**Assumption 2.** (*Instrument Exogeneity*) Marginal costs are  $c_{0jt} = \bar{c}_{0j}(q_{jt}, w_{jt}) + \omega_{0jt}$  for each  $j$ , and  $z_{jt}$  is a vector of  $K$  excluded instruments such that  $E[\omega_{0jt} \mid w_{jt}, z_{jt}] = 0$  *a.s.*<sup>4</sup>

We further simplify the environment with the following assumption:

**Assumption 3.** (*Constant Marginal Cost*) Marginal costs are constant in quantities and only depend on the observable cost shifters  $w_{jt}$ , or  $\bar{c}_{0j}(q_{jt}, w_{jt}) = \bar{c}_{0j}(w_{jt})$  for all  $j$ . Further,  $\bar{c}_{0j}$  is differentiable in  $w_{jt}$ , with partial derivatives  $\frac{\partial \bar{c}_{0j}}{\partial w_{jt}} \neq 0$  everywhere.

For our running example and our empirical application, we will assume that marginal costs are linear in cost shifters,  $\bar{c}_{0j}(w_{jt}) = w_{jt}\gamma$ , as is standard in the empirical literature; but we make no such assumption for our general falsification results.

Berry and Haile (2014) point out that several sources of variation are typically exogenous, which can then be used to construct instruments. These include rival cost shifters and own and rival product characteristics. We refer throughout to *cost side instruments* and *demand side instruments*, respectively, as those formed with this variation.

When true costs and markups are unobserved, the exclusion restriction summarizes what we know about the true model:  $E[\omega_{0jt} \mid w_{jt}, z_{jt}] = 0$  holds for the true cost function. For a candidate model  $m$  and candidate cost function  $\bar{c}_{mj}$ , we can use Equation (1) to define  $\omega_{mjt}$ , a model specific analog of  $\omega_{0jt}$ , as  $\omega_{mjt} = p_{jt} - \Delta_{mjt} - \bar{c}_{mj}(w_{jt})$ . Thus,  $E[\omega_{mjt} \mid w_{jt}, z_{jt}]$  provides a natural way to measure a model's fit. When  $E[\omega_{mjt} \mid w_{jt}, z_{jt}] = 0$ ,<sup>5</sup> the model has the same fit as the true model, so it is not falsified by the instruments  $z_{jt}$ . Conversely, when there are no cost functions  $\{\bar{c}_{mj}(w_{jt})\}_{j=1}^J$  satisfying this falsifiable restriction almost surely over the values of  $w_{jt}$  and  $z_{jt}$ , the model is *falsified by the instruments*  $z_{jt}$ .

While instruments may be exogenous, they need not be relevant for falsification, meaning the falsifiable restriction may not be violated for a model  $m$  that is not the truth. However, determining instrument relevance from the falsifiable restriction is not obvious. To focus on the economic determinants of instrument relevance, we next connect the intuition of the pass-through regression to the instrument-based approach.

## 2.3 Connecting Pass-through Regression to Exclusion Restriction

Consider again the setting of Section 2.1: the true model is in the rule-of-thumb class, we have variation in unit taxes  $\tau_t$ , and have residualized all variables with respect to  $w_{jt}$ . As we

---

<sup>4</sup>Here, as in the rest of the paper, the expectation is taken over realizations of unobservables, and *almost surely (a.s.)* refers to realizations of the exogenous observables  $(w_{jt}, z_{jt})$ .

<sup>5</sup>This expression is the analogue of the condition in Theorem 9 of Berry and Haile (2014).

are making a connection to regression, we implement the falsifiable restriction as a covariance restriction using  $\tau_t$  as an instrument. For models  $m$  in the rule-of-thumb class, where  $p_{jt} = \phi_m c_{jt} = \phi_m (\bar{c}_{jt} + \tau_t + \omega_{jt})$ , we can rewrite  $E[\omega_{mjt}\tau_t] = \phi_m^{-1} E[p_{jt}\tau_t] - E[(\tau_t + \omega_{0jt})\tau_t]$ .<sup>6</sup> By Assumption 2,  $\tau_t$  is uncorrelated with  $\omega_0$ , so that model  $m$  is falsified when  $\phi_m \neq (E[\tau_t\tau_t])^{-1} E[p_{jt}\tau_t]$ . Since  $(E[\tau_t\tau_t])^{-1} E[p_{jt}\tau_t] = \beta_{PTR}$ , falsifying a rule-of-thumb model with unit tax instruments amounts to comparing  $\phi_m$  to the coefficient of the pass-through regression. In this limited setting, the instrument-based approach compares measured pass-through with pass-through implied by model  $m$  and thus coincides with the regression approach.

This connection can be further used to understand instrument relevance. Consider again the example of falsifying a rule-of-thumb model, but now using a rival firm's cost shifter as an excluded instrument, which we denote  $z_{jt}$ . Model  $m$  is not falsified by this instrument, because both  $E[p_{jt}z_{jt}] = E[\tau_t z_{jt}] = 0$ . Economically, in the rule-of-thumb class, pass-through of a rival's cost is zero, so variation in a rival's cost shifters has no effect on market outcomes: prices and shares. Notice that this can be cast as an irrelevance problem for an IV version of the pass-through regression, where  $z_{jt}$  is used to instrument for  $\tau_t$ . In the next section, we will show that differences in pass-through continue to underpin instrument relevance beyond this simple environment where pass-through is summarized by a parameter.

### 3 Role of Pass-Through for Falsification

To gain insight into the economic determinants of instrument relevance, it is useful to follow Backus et al. (2021) and rewrite the falsifiable restriction in terms of markups, an economic implication of the model. As prices in the data are generated from the true model in Equation (1), we can express the model specific unobservable  $\omega_{mjt}$  as  $\omega_{mjt} = \Delta_{0jt} - \Delta_{mjt} + \bar{c}_{0j}(w_{jt}) - \bar{c}_{mj}(w_{jt}) + \omega_{0jt}$ . From this, we can restate the falsifiable restriction as follows.<sup>7</sup>

**Lemma 1.** *Under Assumptions 1-3, model  $m$  is falsified by instruments  $z_{jt}$  if and only if for some  $j$  there exists no function  $\bar{c}_{mj}(w_{jt})$  such that*

$$E[\Delta_{0jt} - \Delta_{mjt} \mid w_{jt}, z_{jt}] = \bar{c}_{mj}(w_{jt}) - \bar{c}_{0j}(w_{jt}) \quad a.s.$$

We want to understand the economic features of a model that underlie the empirical content of this statistical condition and therefore distinguish models of conduct.<sup>8</sup> Since this

<sup>6</sup>Note that residualized variables have zero mean, so for two variables  $x$  and  $y$ ,  $E[xy] = \text{cov}(x, y)$ .

<sup>7</sup>Proofs of all lemmas, propositions, and corollaries are in Appendix A.

<sup>8</sup>Previous work (Backus et al., 2021; Duarte et al., 2024) has explored statistical implications of this condition in the context of the RV model selection test.

condition involves the conditional expectation of a variable with respect to an instrument, a useful step is to consider how the variable moves with a change in the instrument. Thus, it is useful to restate Lemma 1 in terms of the marginal impacts of the instruments on markups. To do this, we will assume markups do not move discontinuously as instruments change:<sup>9</sup>

**Assumption 4.** (*Continuous Markups*) For any model  $m$ , for any  $j$ ,  $E[\Delta_{mjt} \mid \mathbf{w}_{jt}, z_{jt}]$  is absolutely continuous in  $z_{jt}$ .

Based on Lemma 1, for a model not to be falsified, the conditional expectation of  $\Delta_{0jt} - \Delta_{mjt}$  must match the difference in implied cost at each value of the instruments  $z_{jt}$ . As instruments are excluded from cost, a marginal change in any of the  $K$  instruments has no effect on the implied costs for either model. This means that if a model is not falsified, the impact of the instruments on the conditional expectation of  $\Delta_{0jt} - \Delta_{mjt}$  must be zero. We focus on the limit of this difference,

$$\lim_{h \rightarrow 0} \frac{E[\Delta_{0jt} - \Delta_{mjt} \mid \mathbf{w}_{jt}, z_{jt} = \tilde{z}_{jt} + h_k] - E[\Delta_{0jt} - \Delta_{mjt} \mid \mathbf{w}_{jt}, z_{jt} = \tilde{z}_{jt}]}{h}$$

where  $h_k$  is a  $K$ -vector of zeros with the scalar  $h$  in the  $k$ -th position. The marginal effect of the  $k$ -th instrument  $z_{jt}^{(k)}$  on the conditional expectation of  $\Delta_{0jt} - \Delta_{mjt}$  is the average difference in the marginal effect of  $z_{jt}^{(k)}$  on  $\Delta_{0jt}$  and  $\Delta_{mjt}$ , giving the following:

**Lemma 2.** Under Assumptions 1-4, model  $m$  is falsified by instruments  $z_{jt}$  if and only if for some  $j$  and  $k$ ,

$$E \left[ \frac{d\Delta_{0jt}}{dz_{jt}^{(k)}} - \frac{d\Delta_{mjt}}{dz_{jt}^{(k)}} \mid \mathbf{w}_{jt}, z_{jt} \right] \neq 0 \quad w.p.p.$$

We illustrate the intuition behind Lemma 2 in the following example:

*Example 1:* In this and every subsequent example, we consider a simple environment with two single-product firms and logit demand, so that market shares  $j \in \{1, 2\}$  are given by

$$s_{jt} = \frac{\exp(x_{jt}\beta - \alpha p_{jt})}{1 + \exp(x_{1t}\beta - \alpha p_{1t}) + \exp(x_{2t}\beta - \alpha p_{2t})},$$

where  $x_{jt}$  are characteristics of product  $j$  in market  $t$  and  $\alpha$  and  $\beta$  are coefficients. Further suppose that there is no unobservable variation, so that demand (as a function of price) and the cost of firm 1,  $c_{01t}$ , remain fixed.

---

<sup>9</sup> Assumption 4 holds for all models in this paper, and for all standard models.

First, suppose that the true model generating the data is Bertrand-Nash competition in prices, which we will hereafter refer to as Bertrand, and we want to use rival cost shifters to falsify the model of Keystone pricing, under which  $p_{jt} = 2c_{mjt}$  and  $\Delta_{mjt} = c_{mjt}$ . Suppose that we observe variation in the cost shifter of firm 2,  $w_{2t}$ , so that  $z_{1t} = w_{2t}$ . Under Lemma 2, Keystone is falsified if the instruments differentially move the Keystone and Bertrand markups for firm 1. Under the true Bertrand model, when  $w_{2t}$  increases, both firms' equilibrium prices increase, and since  $c_{01t}$  did not change,  $\Delta_{01t}$  increases by the same amount as  $p_{1t}$ .<sup>10</sup> Under the Keystone model, however, firm 1's equilibrium price does not change with variation in  $w_{2t}$ , all else equal. To rationalize the increase in  $p_{1t}$ , the Keystone model would impute an increase in firm 1's marginal cost  $c_{m1t}$  and markups  $\Delta_{m1t}$  by *half* the amount of the change in  $p_{1t}$ . The difference between  $\frac{d\Delta_{01t}}{dw_{2t}}$  and  $\frac{d\Delta_{m1t}}{dw_{2t}}$  is what allows us to falsify Keystone.●

The derivative of markups with respect to the instruments is an object whose properties depend on the economics of the firm conduct model, so this lemma allows us to connect the econometric perspective on falsifiability with a more theoretical view. To falsify a model, variation in the instruments must induce differential changes in the implied markups for that model and the true model. However, since demand and cost shocks are unobserved, it is not possible to isolate changes solely caused by the instruments; we therefore need the marginal impact of the instruments to differ across the two models when we average over the unobserved shocks.

Even beyond the restricted settings considered in Section 2.1 where the true own cost pass-through can be consistently estimated via regression, it is still largely the cost pass-through of a model that determines falsifiability. To see this, note that the markups for the true data generating process and the model to be falsified are functions of two endogenous variables, market shares and prices. As the demand system makes market shares a function of prices, we can write the vector  $\Delta_{mt}$  as a function of prices, instruments and other exogenous variables,  $\Delta_{mt} = \Delta_m(p_t, z_t, w_t, x_t, \omega_{mt}, \xi_t)$ .<sup>11</sup>

Instruments may affect markups either directly, or through prices. In light of Lemma 2, we are interested in the average difference between the causal effects of an instrument  $z_{jt}^{(k)}$

---

<sup>10</sup>In Example 2, we will derive the cost pass-through matrix for the Bertrand model, and see that it has positive off-diagonal terms. The economic intuition is that when firm 2 raises its price in response to an increase in costs, this decreases the own-price elasticity of firm 1, leading firm 1 to raise its price as well.

<sup>11</sup>This holds even if markups depend directly on marginal costs as in Keystone pricing, since marginal costs are themselves determined by  $w_{jt}$  and  $\omega_{mjt}$ .

on model  $m$  and the truth. Letting  $(A)_j$  denote the  $j^{\text{th}}$  row of a vector or matrix, we have

$$\frac{d\Delta_{0jt}}{dz_{jt}^{(k)}} - \frac{d\Delta_{mjt}}{dz_{jt}^{(k)}} = \underbrace{\frac{\partial\Delta_{0jt}}{\partial z_{jt}^{(k)}} - \frac{\partial\Delta_{mjt}}{\partial z_{jt}^{(k)}}}_{\text{Direct Effect}} + \underbrace{\left( \frac{\partial\Delta_{0t}}{\partial p_t} - \frac{\partial\Delta_{mt}}{\partial p_t} \right)_j \frac{dp_0}{dz_{jt}^{(k)}}}_{\text{Indirect Effect}}$$

where  $p_0(\cdot)$  is the function mapping primitives to equilibrium prices under the true model. From this expression we see two distinct effects of instruments on the difference in markups. The first is the *direct effect* of instruments, or  $\frac{\partial\Delta_{0jt}}{\partial z_{jt}^{(k)}} - \frac{\partial\Delta_{mjt}}{\partial z_{jt}^{(k)}}$ . This element is non-zero whenever instruments such as product characteristics differentially enter the markup functions of model  $m$  and the truth. The second term represents the *indirect effect*, which happens through prices. We further investigate the economic content of this term. For expositional clarity, we make a simplifying assumption:

**Assumption 5.** (*No Direct Effect of Costs*) The markup functions  $\Delta_{0t}$  and  $\Delta_{mt}$  for both the true model and the model to be falsified do not depend directly on marginal costs.

Most standard conduct models based on profit maximization satisfy this assumption. Moreover, as long as the mapping from marginal costs to prices is invertible, any dependence of markups on costs can be restated as a dependence on prices to satisfy Assumption 5. For example, under Keystone pricing  $\Delta_{Kjt} = c_{jt}$ , but we can alternatively write  $\Delta_{Kjt} = \frac{1}{2}p_{jt}$  to remove the direct dependence on costs and satisfy Assumption 5, which therefore serves more as a labeling convention than a substantive assumption.<sup>12</sup>

For any model  $m$  we can compute the cost pass-through matrix  $P_{mt}$  via the Implicit Function Theorem. Writing the first-order conditions for any model  $m$  in market  $t$  as

$$F_m(p_t, c_t) = p_t - c_t - \Delta_{mt} = 0$$

equilibrium prices under model  $m$  are an implicit function of costs,  $p_t = p_m(c_{mt})$ , defined as the solution to  $F_m(p_m(c_{mt}), c_{mt}) = 0$ . Thus, for any model  $m$ ,  $p_m$  is defined implicitly by these first-order conditions; under Assumption 5, the Implicit Function Theorem gives

$$P_{mt} = \frac{dp_m(\cdot)}{dc_t} = - \left[ \frac{\partial F_m}{\partial p_t} \right]^{-1} \frac{\partial F_m}{\partial c_t} = (I - H_{\Delta_{mt}})^{-1}$$

where  $H_{\Delta_{mt}} = \frac{\partial\Delta_{mt}}{\partial p_t}$ . In the indirect effect, we can rewrite the difference in price derivatives of markups in terms of inverse pass-through matrices, because  $P_{mt}^{-1} = \left( I - \frac{\partial\Delta_{mt}}{\partial p_t} \right)$  as long as

---

<sup>12</sup>Assumption 5 can fail when the mapping from costs to prices is not invertible – for example, if a firm priced a product at \$50 whenever marginal costs were between \$30 and \$40.

$P_{mt}$  is invertible. We make the assumption that it is:<sup>13</sup>

**Assumption 6.** (*Invertibility of Pass-throughs*) For any model  $m$ , in any market  $t$ , the pass-through matrix  $P_{mt}$  has full rank.

*Example 2:* We return to the “example environment” of two single-product firms under simple logit demand, and calculate equilibrium markups and pass-through matrices for four models we will use extensively for intuition. First is marginal cost pricing,  $p_{jt} = c_{0jt}$ , under which markups are  $\Delta_{MCjt} = 0$ , and pass-through and inverse pass-through matrices are  $P_{MCjt} = P_{MCjt}^{-1} = I$ . Second is the Keystone model, where prices are set at twice marginal costs; this gives markups  $\Delta_{Kt} = \frac{1}{2}p_t$ , cost pass-through matrix  $P_{Kt} = 2I$ , and inverse pass-through matrix  $P_{Kt}^{-1} = \frac{1}{2}I$ . Next is Bertrand price competition, which in this setting gives markups, inverse pass-through, and pass-through matrices

$$\Delta_{Bt} = \begin{bmatrix} \frac{1}{\alpha(1-s_{1t})} \\ \frac{1}{\alpha(1-s_{2t})} \end{bmatrix}, \quad P_{Bt}^{-1} = \begin{bmatrix} \frac{1}{1-s_{1t}} & -\frac{s_{1t}s_{2t}}{(1-s_{1t})^2} \\ -\frac{s_{1t}s_{2t}}{(1-s_{2t})^2} & \frac{1}{1-s_{2t}} \end{bmatrix}, \quad \text{and} \quad P_{Bt} = \kappa_{Bt} \begin{bmatrix} \frac{1}{1-s_{2t}} & \frac{s_{1t}s_{2t}}{(1-s_{1t})^2} \\ \frac{s_{1t}s_{2t}}{(1-s_{2t})^2} & \frac{1}{1-s_{1t}} \end{bmatrix}$$

where  $s_{0t} = 1 - s_{1t} - s_{2t}$  and  $\kappa_{Bt} = s_{0t}^{-1}(1 - s_{1t})^2(1 - s_{2t})^2$ . Finally, there is the differentiated-products version of Cournot competition in quantities, where each firm chooses a market share  $s_{jt}$  (taking the other firm’s share as given) to maximize profits. This gives markups, inverse pass-through, and pass-through

$$\Delta_{Ct} = \begin{bmatrix} \frac{1-s_{2t}}{\alpha s_{0t}} \\ \frac{1-s_{1t}}{\alpha s_{0t}} \end{bmatrix}, \quad P_{Ct}^{-1} = \begin{bmatrix} \frac{1-s_{2t}}{s_{0t}} & 0 \\ 0 & \frac{1-s_{1t}}{s_{0t}} \end{bmatrix}, \quad \text{and} \quad P_{Ct} = \begin{bmatrix} \frac{s_{0t}}{1-s_{2t}} & 0 \\ 0 & \frac{s_{0t}}{1-s_{1t}} \end{bmatrix}$$

Assumption 6 holds for all four conduct models. Also in this simple environment, the Cournot model has zero pass-through of rival’s costs (or zero off-diagonal elements for  $P_{Ct}$ ) like marginal cost pricing and Keystone pricing. This will be important in future examples. •

Breaking the terms in Lemma 2 into the direct and indirect effects yields the following:

---

<sup>13</sup>This assumption has economic content, as it requires that each product has non-zero pass-through for at least one cost (either own or rival) in the market. Moreover, pass-through vectors for each product cannot be linear combinations of those of other products, meaning that costs must affect different products in a distinct way. This is satisfied in most models, as the pass-through of own cost that is measured by the main diagonal of  $P_{mt}$  is typically different than the pass-through of rival costs.

**Proposition 1.** *Under Assumptions 1-6, a model  $m$  is falsified by instruments  $z_{jt}$  if and only if for some  $j$  and  $k$ ,*

$$E \left[ \frac{\partial \Delta_{0jt}}{\partial z_{jt}^{(k)}} - \frac{\partial \Delta_{mjt}}{\partial z_{jt}^{(k)}} + (P_{mt}^{-1} - P_{0t}^{-1})_j \frac{dp_0}{dz_{jt}^{(k)}} \mid \mathbf{w}_{jt}, z_{jt} \right] \neq 0 \quad w.p.p. \quad (2)$$

This proposition casts falsification in terms of the direct and indirect effect of the instruments. When the instruments move the markup functions directly and differentially, this will likely enable falsification. Furthermore, if instruments indirectly affect markups through prices, falsification depends on  $(P_{mt}^{-1} - P_{0t}^{-1})$ , the difference in inverse pass-throughs of the true model and model  $m$ . Similar to the intuition obtained from the pass-through regression, differences in inverse pass-throughs (and therefore pass-throughs) permit falsification, though only if instruments move prices under the true model, i.e.  $\frac{dp_0}{dz_{jt}^{(k)}} \neq 0$ .

The general framework we develop has practical applications. We highlight two in the next sections: evaluating the relevance of specific instruments, and motivating instrument selection when a researcher's objective is to perform a particular counterfactual.

## 4 Evaluating Instrument Relevance

Next, we evaluate the relevance of instruments commonly used in the literature. We start with standard cost side instruments (based on rival cost shifters) and demand side instruments (based on product characteristics). We then augment Proposition 1 for settings with taxes, and show the relevance of tax instruments.

### 4.1 Cost Side Instruments

The implications of Proposition 1 are particularly stark when the instruments are formed with rival cost shifters. Under Assumption 5, the instruments have no direct effect on markups, and enter either markup function only through  $p_0(\cdot)$ . Note that since the marginal cost function  $\bar{c}_{mj}$  is not observed directly, a firm's own cost shifters are irrelevant instruments for conduct: for any model  $m$ , we can always specify a cost function  $\bar{c}_{mj}$  giving  $E[\omega_{mjt} | \mathbf{w}_{jt}] = 0$ .<sup>14</sup> Thus, only observable shifters of a *rival's* cost are plausibly relevant instruments.

*Example 3:* We return to the example environment. Suppose we want to falsify Bertrand competition when the true model is Cournot competition, using variation in the cost shifter of firm 2,  $\mathbf{w}_{2t}$ , so that  $z_{1t} = \mathbf{w}_{2t}$ . Assume  $\mathbf{w}_{2t}$  is scalar, and  $\bar{c}_{02}(\mathbf{w}_{2t}) = \gamma \mathbf{w}_{2t}$ . Under the true

---

<sup>14</sup>For any model  $m$ , define  $\bar{c}_{mj}(\mathbf{w}_{jt}) = E[p_{jt} | \mathbf{w}_{jt}] - E[\Delta_{mjt} | \mathbf{w}_{jt}]$ , and  $E[\omega_{mjt} | \mathbf{w}_{jt}] = 0$  follows.

model, variation in  $w_{2t}$  only moves the price of firm 2, with  $\frac{dp_0}{dz_{1t}} = \left[ \gamma \frac{0}{1-s_{1t}} \right]$ . In this simplified environment where there is no direct effect of the instruments under either model and where the unobservables are held fixed, falsification under Proposition 1 obtains if in any market  $t$ ,  $(P_{Bt}^{-1} - P_{Ct}^{-1})_1 \left[ \gamma \frac{0}{1-s_{1t}} \right] \neq 0$  or  $\gamma \frac{s_{0t}}{1-s_{1t}} ((P_{Bt}^{-1})_{12} - (P_{Ct}^{-1})_{12}) \neq 0$ . Thus, Bertrand is falsified by variation in  $w_{2t}$ , as the (1,2)-elements of  $P_{Bt}^{-1}$  and  $P_{Ct}^{-1}$  differ. If instead we wished to falsify either the Keystone model or marginal cost pricing, we would fail: pass-through matrices under both Cournot and the model to be falsified are diagonal, so  $(P_{Ct}^{-1})_{12} = (P_{Kt}^{-1})_{12} = (P_{MCt}^{-1})_{12} = 0$ , and rival cost shifters are irrelevant for falsifying either of the two models. •

To move from Example 3 to our general environment, define the  $J \times J$  matrix  $[P_m^{-1}P_0]^*$  by

$$[P_m^{-1}P_0]_j^* = E [P_{mt}^{-1}P_{0t} \mid w_{jt}, z_{jt}]_j.$$

That is,  $[P_m^{-1}P_0]^*$  is the expected value of  $P_{mt}^{-1}P_{0t}$ , but where the expectation in the  $j$ -th row is taken conditional on the realized values of product  $j$ 's cost shifters and instruments  $(w_{jt}, z_{jt})$ . Thus,  $[P_m^{-1}P_0]^*$  is a function of the full vector of observables  $(w_t, z_t)$ , with each row depending on a different subset of those observables. This matrix arises from Equation (2) because  $\frac{dp_0}{dz_{jt}^{(k)}}$  will often be proportional to  $P_{0t}$ , so  $(P_{mt}^{-1} - P_{0t}^{-1}) \frac{dp_0}{dz_{jt}^{(k)}} \propto (P_{mt}^{-1}P_{0t} - I)$ .

Given observable and unobservable variation, falsifiability depends on the matrix  $[P_m^{-1}P_0]^*$ :

**Corollary 1.** *Suppose that for each product  $j$ , the vector of instruments  $z_{jt}$  includes a cost shifter of every rival product. Under Assumptions 1-6, model  $m$  is falsified by the instruments if with positive probability over  $(w_t, z_t)$ , the matrix  $[P_m^{-1}P_0]^*$  is not diagonal.*

For intuition, put aside the expectation over unobservables, and focus on model  $m$  being falsified by rival cost shifters if  $P_m^{-1}P_0$  is not diagonal. This is when  $P_0 \neq P_m D$  for any diagonal matrix  $D$ , i.e., when some column of  $P_m$  is not a scalar multiple of the corresponding column of  $P_0$ . Assuming own cost pass-through rates are nonzero, this in turn happens if and only if there is some  $(j, \ell)$  such that the ratio

$$\frac{dp_{\ell t}}{dc_{jt}} \bigg/ \frac{dp_{jt}}{dc_{jt}}$$

differs between model  $m$  and the true model. If instead this ratio is the same across the two models, then model  $m$  can be rationalized by the unknown  $\frac{\partial \bar{c}_{mj}}{\partial w_{jt}}$ , leading to lack of falsification with cost side instruments.<sup>15</sup> For example, if  $P_{mt}$  and  $P_{0t}$  are always both diagonal matrices, then this ratio is always zero and model  $m$  is not falsified by cost side instruments. Thus,

<sup>15</sup>Suppose a shifter of cost  $j$  moves, resulting in changes to prices  $p_j$  and (potentially)  $p_\ell$ . The change in  $p_j$  is proportional to  $\frac{dp_{0jt}}{dc_{jt}} \frac{\partial \bar{c}_{0j}}{\partial w_{jt}}$ ; the change in  $p_\ell$  is proportional to  $\frac{dp_{0\ell t}}{dc_{jt}} \frac{\partial \bar{c}_{0j}}{\partial w_{jt}}$ . The effect of the cost shifter

we can think of cost side instruments targeting the ratio of a rival product's pass-through of a product  $j$ 's cost to product  $j$ 's pass-through of its own cost.<sup>16</sup>

## 4.2 Demand Side Instruments

**The General Case:** In the general case, Proposition 1 characterizes when a model can be falsified using product characteristics. Two key differences arise between the case of cost side instruments just discussed and the case of demand side instruments. The first is that product characteristics are sometimes excluded from cost, allowing the researcher to use variation in both own and rival products' characteristics as valid instruments. This feature makes the indirect effect of instruments more useful: insofar as product characteristics are excluded from marginal cost, the ability to use own product characteristics as instruments will help the researcher trace out the full pass-through matrix, as opposed to relying on only the ratio of off-diagonal to diagonal elements for falsification. Thus, models with different diagonal pass-through matrices could still be falsified with demand side instruments. Second, in most standard conduct models, product characteristics will directly enter the markup function and therefore have a direct effect on the implied costs, so that  $\frac{\partial \Delta_{mjt}}{\partial z_{jt}^{(k)}} \neq 0$ . Adding an extra channel in general makes falsification easier, although there may be cases in which direct effects cancel out indirect effects, thus hindering falsification.

**An Interesting Special Case:** For a subset of conduct models and demand systems, we can simplify Equation (2) to a condition that is easier to interpret. For this result, we focus on models that satisfy two conditions. First, the first-order conditions (1) must be a system of equations containing only markups, market shares, and the derivatives of market shares with respect to prices; all other variables only enter into the first-order conditions through these. Second, demand must have a linear index structure. We state these conditions as formal assumptions next, beginning with the former.

**Assumption 7.** (*No Direct Effects of Primitives*) For both the true model and the model to be falsified, the first-order condition can be written as a function of only markups, market shares, and price derivatives of demand  $\frac{\partial s_t}{\partial p_t}$ .

Under Assumption 7,  $\Delta_{mt} = \Delta_m \left( s_t, \frac{\partial s_t}{\partial p_t} \right)$  so that the implied markup function for any model  $m$  only depends directly on  $s_t$  and  $\frac{\partial s_t}{\partial p_t}$ . In Appendix C, we demonstrate a range

---

$\frac{\partial \bar{c}_{0j}}{\partial w_{jt}}$  is not known ex ante; but for a model where  $\frac{dp_{m\ell t}}{dc_{jt}} / \frac{dp_{mjt}}{dc_{jt}}$  differs from the truth, no value of  $\frac{\partial \bar{c}_{mj}}{\partial w_{jt}}$  can rationalize both price changes, so  $\omega_{mjt}$  can't be mean-independent of  $w_{jt}$ .

<sup>16</sup>Falsifiability does not require every product's instruments  $z_{jt}$  to include a shifter of every rival product's cost, just that there be some pair  $(j, \ell)$ ,  $\ell \neq j$ , where  $[P_m^{-1} P_0]_{j, \ell}^* \neq 0$  and  $z_{jt}$  includes a shifter of  $c_{0\ell}$ .

of conduct models which satisfy this assumption. These include all models where firms act to maximize their profits under some assumption about how rival firms will respond to their choices, such as Bertrand, Cournot, and Stackelberg; they also includes many collusive models, and models of altruistic firms who maximize a weighted sum of own profits and consumer surplus. For example, under Bertrand competition the implied markups are

$$\Delta_{Bt} = - \left[ \Omega \odot \left[ \frac{\partial s_t}{\partial p_t} \right]' \right]^{-1} s_t,$$

where  $\Omega$  is the ownership matrix, with  $\Omega_{ij} = 1$  if products  $i$  and  $j$  are sold by the same firm and zero otherwise. Here, we see that all primitives enter the equilibrium markup only through their effects on  $\left( s_t, \frac{\partial s_t}{\partial p_t} \right)$ . The assumption is also satisfied by marginal cost pricing, where  $\Delta_{MCt} = 0$ . However, the assumption fails in Keystone pricing, where  $p_t$  enters the implied markups directly:  $\Delta_{Kt} = \frac{1}{2}p_t$ . The assumption also fails in models where firms' objective functions depend in part on revenue rather than profit.

Under an additional assumption on the demand system, the effect of product characteristics becomes very similar to the effect of cost shifters on markups:

**Assumption 8.** (*Demand Index*) Demand depends on  $x_t$  and  $p_t$  only through a one-dimensional index  $\delta_t = x_t\beta - \alpha p_t + \xi_t$ , with  $\alpha > 0$ , all elements of  $\beta$  nonzero, and  $s_t = s(x_t, p_t, \xi_t, \cdot) = s(\delta_t, \cdot)$ .

This assumption is satisfied when the demand system is logit or nested logit; it is not generally satisfied for mixed logit demand models with random coefficients on either price or characteristics, but we expect the added variation in richer models to make falsification easier, not harder.<sup>17</sup>

Together, Assumptions 7 and 8 imply that the markup functions  $\Delta_0$  and  $\Delta_m$  depend on prices and product characteristics only through  $\delta_t$ . This, in turn, implies that  $x_t$  and  $c_t$  affect equilibrium markups only through the term  $x_t\beta - \alpha c_t$ , and therefore that the effects of marginal costs and product characteristics on markups are virtually identical. For intuition, consider the Bertrand model with single-product firms. Instead of choosing prices, we can think of firms directly choosing markups  $\Delta_{jt}$  to maximize  $(p_{jt} - c_{jt})s_j(\delta_{jt}, \delta_{-jt}) = \Delta_{jt}s_j(x_{jt}\beta - \alpha c_{jt} + \xi_{jt} - \alpha\Delta_{jt}, \delta_{-jt})$ . The optimal  $\Delta_{jt}$  depends on market primitives only through the term  $x_{jt}\beta - \alpha c_{jt} + \xi_{jt}$ . The same logic extends easily to all the models discussed in Appendix C.

---

<sup>17</sup>Alternatively, we can define  $\delta_t = x_t\beta - p_t$  and  $s_t = s(x_t, p_t, \xi_t, \cdot) = s(\delta_t, \xi_t, \cdot)$ . The proceeding results still apply, and this admits mixed logit models as long as some product characteristic is a perfect substitute for income.

*Example 4:* In Example 3, cost side instruments could not falsify Keystone or marginal cost pricing when the truth is Cournot. Here, we consider falsification with demand side instruments. Let  $x_{jt}$  be a scalar characteristic of product  $j$ , and suppose  $z_{1t} = x_{1t}$ .

First, consider falsifying marginal cost pricing. Our example environment satisfies Assumption 8, and both the Cournot model and marginal cost pricing satisfy Assumption 7; so for both models,  $x_{1t}$  and  $c_{1t}$  affect firm 1's markups only through the term  $x_{1t}\beta - \alpha c_{1t}$ , and therefore  $\frac{d\Delta_{1t}}{dx_{1t}} = -\frac{\beta}{\alpha} \frac{d\Delta_{1t}}{dc_{1t}}$ . Under Lemma 2, then, the model of marginal cost pricing will be falsified if and only if  $\frac{d\Delta_{MC1t}}{dc_{1t}} \neq \frac{d\Delta_{C1t}}{dc_{1t}}$ . Indeed, under marginal cost pricing, markups are zero, and therefore  $\frac{d\Delta_{MC1t}}{dc_{1t}} = 0$ ; while under Cournot,  $\frac{d\Delta_{C1t}}{dc_{1t}} = [P_{Ct}]_{11} - 1 = -\frac{s_1}{s_0 + s_1} \neq 0$ ; so the marginal cost pricing model is falsified.

Falsifying Keystone pricing is subtler, because markups  $\Delta_{Kjt} = 0.5p_{jt}$  violate Assumption 7. However, since product characteristics do not move prices, demand side instruments have neither a direct nor an indirect effect on markups, and  $\frac{d\Delta_{K1t}}{dx_{1t}} = 0$ . As noted above, under Cournot,  $\frac{d\Delta_{C1t}}{dx_{1t}} = -\frac{\beta}{\alpha} \frac{d\Delta_{C1t}}{dc_{1t}} = \frac{\beta}{\alpha} \frac{s_1}{s_0 + s_1} \neq 0$ , so the Keystone model is falsified as well.

Finally, if the truth were Keystone, demand side instruments would be irrelevant for falsifying marginal cost pricing, and vice versa, since  $\frac{d\Delta_{m1t}}{dx_{1t}} = 0$  for both models. •

Under Assumptions 7 and 8, demand side instruments shift markups in the same way as marginal costs, suggesting that instrument relevance will be similar to cost side instruments. However, while own cost shifters are not excluded from costs (and therefore can't be used to show a violation of the exclusion restriction), own product characteristics *are* excluded. Whereas falsification using cost shifters depends on the *off-diagonal* terms of  $P_m^{-1}P_0$ , falsification using demand side instruments can use the diagonal terms as well. We can formalize this as follows:

**Corollary 2.** *Suppose that for every product  $j$ , the vector of instruments  $z_{jt}$  includes a product characteristic of every product. Under Assumptions 1-8, model  $m$  is falsified by the instruments  $z_t$  if with positive probability over  $(w_t, x_t)$ , the matrix  $[P_m^{-1}P_0]^*$  is not equal to the identity matrix.*

The intuition is that the marginal effect of a product characteristic  $x_{\ell t}^{(k)}$  on the inferred costs of product  $j$  under model  $m$  is proportional to  $(I - P_{mt}^{-1}P_{0t})_{j\ell}$ . Thus, if this term is nonzero in expectation for some values of observables, a change in  $x_{\ell t}^{(k)}$  would lead to changes in the mean of  $c_{mjt}$ ; since product characteristics used as instruments are excluded from costs, this would again violate the restriction that  $\omega_{mjt}$  is mean-independent of the instruments.

Loosely, while falsifying a model with cost side instruments requires that  $P_{mt}^{-1}P_{0t}$  is not diagonal on average, falsifying a model with demand side instruments only requires that  $P_{mt}^{-1}P_{0t} \neq I$ . Thus, when  $\alpha$  and  $\beta$  are both non-zero, it is easier to falsify model  $m$  using

product characteristics: under Assumptions 1-8 any model which is falsified by rival cost shifters is also falsified by product characteristics. This is because under Assumptions 7 and 8, the effect of product characteristics on markups is nearly identical to the effect of marginal costs, but scaled by demand system parameters  $\beta$  and  $\alpha$  which are already known rather than through a cost function  $\bar{c}_{mj}(\cdot)$  which is ex ante unknown. This allows demand side instruments to target the entire pass-through matrix, not just the ratio of off-diagonal to diagonal elements, allowing falsification whenever (loosely)  $P_m \neq P_0$ . In the more general case where either Assumption 7 or 8 does not hold, however, demand side instruments target a mix of pass-through and the direct effect of the instrument; this added effect likely helps in falsifying incorrect models, but leaves the source of this falsification more opaque.

### 4.3 Tax Instruments

Next, we consider tax rates as another source of exogenous variation. Suppose that governments levy market-level unit and/or ad valorem taxes, at rates  $\tau_t$  and  $\nu_t$ , respectively.<sup>18</sup> For ad valorem taxes, we will work with  $\nu_t$ , the fraction of the consumer's payment received by the firm, which is  $1 - \nu_t$  when the tax is levied on the firm and  $1/(1 + \nu_t)$  when it is levied on the consumer. For unit taxes, we will assume they are levied on firms.

In the presence of taxes, the after-tax version of the first order condition (1) is

$$\nu_t p_{jt} - \tau_t = \nu_t \Delta_{0jt} + c_{0jt}, \quad (3)$$

where  $\nu_t p_{jt} - \tau_t$  is the after-tax revenue received by the firm for product  $j$  in market  $t$ .

Similar to Assumption 5, we assume that taxes do not directly enter markups:

**Assumption 9.** (*No Abnormal Effects of Tax Rates*) The markup functions  $\Delta_{0t}$  and  $\Delta_{mt}$  for the true model and the model to be falsified do not depend directly on tax rates.

This can be thought of as an assumption that taxes are fully salient and there is no avoidance/evasion. For example, under the Keystone model markups are  $\Delta_{Kjt} = \frac{1}{2}p_{jt}$ . In the presence of a unit tax and no ad valorem tax, this means  $p_{Kjt} = 2(\tau_t + c_{Kjt})$ ; the firm perceives the unit tax as part of its marginal costs. With an ad valorem tax instead,  $p_{Kjt} = 2c_{Kjt}/\nu_t$ ; the firm sets price such that after-tax revenue is twice marginal cost.

If variation in taxes is available, and assumed to be exogenous, we can use it to construct *tax instruments*. We can rewrite  $\omega_{mt}$  as  $\omega_{mjt} = \nu_t \Delta_{0jt} - \nu_t \Delta_{mjt} + \bar{c}_{0j}(w_{jt}) - \bar{c}_{mj}(w_{jt}) + \omega_{0jt}$ . Falsification depends on the impossibility of finding a cost function  $\bar{c}_{mj}$  giving  $E[\omega_{mjt}|w_{jt}, z_t] = 0$ ; since  $E[\omega_{0jt}|z_t] = 0$  and the instruments are excluded from  $\bar{c}_{0j}$  and  $\bar{c}_{mj}$ ,

---

<sup>18</sup>Note that taxes apply uniformly to all products, so  $\tau_t$  and  $\nu_t$  are scalars, not vectors.

falsification occurs if  $E[\nu_t(\Delta_{0jt} - \Delta_{mjt}) | w_{jt}, z_t]$  varies with the instrument. Under Assumption 9,  $\frac{\partial \Delta_{mjt}}{\partial z_t} = \frac{\partial \Delta_{0jt}}{\partial z_t} = 0$  for either type of tax instrument; this lets us restate the falsifiable restriction as follows.

**Proposition 2.** *Under Assumptions 1-6 and 9, a model  $m$  is falsified by the tax instrument  $z_t$  if and only if for some  $j$ ,*

$$E \left[ \nu_t (P_{mt}^{-1} - P_{0t}^{-1})_j \frac{dp_0}{dz_t} + (\Delta_{0jt} - \Delta_{mjt}) \frac{d\nu_t}{dz_t} \mid w_{jt}, z_{jt} \right] \neq 0 \quad w.p.p. \quad (4)$$

That is, after ruling out any direct effect of the instrument entering the markup function, the marginal impact of an instrument on the across-model difference in (after-tax) markups still has two potential effects: the indirect effect through prices, and a new effect if the instrument is correlated with the ad valorem tax rate. We now discuss the relevance of unit taxes and ad valorem taxes in turn.

**Unit Tax Instruments:** We first consider unit tax instruments. For each model  $m$ , there is an implied “composite” marginal cost  $\tilde{c}_{mjt} = \bar{c}_{mt}(w_{jt}) + \tau_t + \omega_{mjt}$ , such that  $\nu_t(p_{jt} - \Delta_{mjt}) = \tilde{c}_{mjt}$ . Therefore, marginal cost is composed of two types of observed cost shifters:  $w_{jt}$  which enters in an ex ante unknown way, and  $\tau_t$  which enters additively with a known coefficient of 1. Since this coefficient is known ex ante, variation in a firm’s *own* marginal cost induced by the unit tax can serve as exogenous variation for distinguishing conduct.

We noted in Example 4 that in our example environment, traditional demand side instruments fail to falsify a model of marginal cost pricing when the true model is Keystone pricing. Further, since marginal cost pricing and Keystone both have diagonal pass-through matrices, instruments based on rival cost shifters would also fail to falsify marginal cost pricing under Keystone truth. Here, we illustrate that unit tax instruments do allow falsification.

*Example 5:* Fix  $\nu_t = 1$  (no ad valorem tax), and suppose that we observe variation in a unit tax which is levied on both firms,  $\tau_t$ , so that  $z_t = \tau_t$ . Further suppose we are in a simplified environment where the unobservables and other cost shifters are held fixed. Under Keystone,  $\Delta_{0jt} = \frac{1}{2}p_{jt}$ , and under marginal cost pricing,  $\Delta_{mjt} = 0$ , so the tax rate  $\tau_t$  does not enter directly into either markup function. Further,  $P_{0t} = 2I$  and  $P_{mt} = I$ , and under the true model,  $\frac{dp_0}{d\tau_t} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ , since unit variation in  $\tau_t$  moves both firms’ costs by 1, increasing both prices by 2 under the true model (Keystone). Under Proposition 2, then, falsification occurs because  $(P_{mt}^{-1} - P_{0t}^{-1})_1 \frac{dp_0}{d\tau_t} = ([1 \ 0] - [\frac{1}{2} \ 0]) \begin{bmatrix} 2 \\ 2 \end{bmatrix} \neq 0$ . •

In our general environment, we can establish the following:

**Corollary 3.** *Suppose that for each product  $j$ , the instrument  $z_t$  is the unit tax  $\tau_t$ . Under Assumptions 1-6 and 9, model  $m$  is falsified by the instrument if with positive probability over  $(w_t, z_t)$ , the matrix  $[P_m^{-1}P_0]^*$  has some row whose elements do not sum to 1.*

While the tax rate acts as a cost shifter, there are two important features that make this result different from Corollary 1. First, because the unit tax shifts marginal costs additively, the variation in unit tax rates is not needed to estimate the effect of  $\tau_t$  on marginal cost, and can therefore be used as instrumental variation. In contrast with other cost shifters, this makes the instrument relevant for falsification even when the true model and the model to be falsified both have diagonal pass-through matrices. However, the unit tax applies to all products equally, rather than shifting marginal costs of one product at a time. For this reason, there may be non-diagonal pass-through matrices for which unit tax instruments do not permit falsification. For example, take  $P_{0t} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$  and  $P_{mt} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ : under either of these two conduct models, a unit increase in the tax rate increases both products' prices by 1, making them observationally equivalent.<sup>19</sup>

**Ad Valorem Tax Instruments:** Next, we consider the relevance of ad valorem tax instruments. Inspection of Equation (3) shows that while unit taxes have an additive effect on markups (or costs), ad valorem taxes instead have a multiplicative effect on markups. This allows variation in the tax rate to distinguish models with different average *levels* of markups, not only those models with different pass-through matrices. Variation in ad valorem tax rates therefore allows falsification in some cases where it would be impossible with any of the other instruments considered so far.

*Example 6:* Return to the example environment, and suppose the true model is one in which pre-tax markups are set to a constant level  $\zeta_0$  across markets, so that  $\Delta_{0jt} = \zeta_{0j}$ . We consider falsifying a model in which firms set constant markups  $\Delta_{mt} = \zeta_m \neq \zeta_0$ . Now,  $P_{0t} = P_{mt} = I$ . Variation in product characteristics, rival cost shifters, or unit taxes have no effect on markups, making cost side, demand side and unit tax instruments irrelevant for falsification. However, when using ad valorem tax instruments, the falsification condition in Proposition 2 becomes  $E[\zeta_m - \zeta_0 \mid w_t, \nu_t] \neq 0$ , which holds for any  $\zeta_m \neq \zeta_0$ . Thus, ad valorem taxes are relevant instruments for distinguishing constant markup models. •

This example generalizes to the following corollary:

---

<sup>19</sup>If the government levied product-specific unit taxes, then falsification would obtain, as  $P_{mt} \neq P_{0t}$ .

**Corollary 4.** *Suppose that for each product  $j$ , the instrument  $z_{jt}$  is  $\nu_t$ , a transformation of the ad valorem tax rate. Let Assumptions 1-6 and 9 hold. Further suppose that either  $P_{mt} = P_{0t}$  or the instrument does not move prices under the true model. Model  $m$  is still falsified as long as  $E[\Delta_{mt} | \mathbf{w}_t, \nu_t] \neq E[\Delta_{0t} | \mathbf{w}_t, \nu_t]$  w.p.p. over  $(\mathbf{w}_t, \nu_t)$ .*

Note that if  $P_{mt} \neq P_{0t}$  and the tax rate does move equilibrium prices, then the two terms in Equation (4) are both nonzero, and could theoretically cancel each other out. In fact, this happens for all rule-of-thumb models where after-tax revenue is set to a fixed multiple of (non-tax) marginal costs: if  $\nu_t p_{jt} = \phi_m c_{mjt}$ , then  $\nu_t \Delta_{mjt} = (\phi_m - 1)c_{mjt}$  does not change with  $\nu_t$ , making falsification of the “wrong” model within this class impossible. For models outside this rule-of-thumb class, the two effects are unlikely to cancel: for example, when comparing Bertrand and Cournot competition, the model with higher markups (Cournot, see Magnolfi, Quint, Sullivan, and Waldfogel, 2022b) is typically associated with lower cost pass-through, which gives the two terms in Equation (4) the same sign, ensuring falsification of the wrong model.

Corollaries 3 and 4, and the discussion above, illustrate the usefulness of tax rates as instruments for conduct. Tax rates enter the equilibrium first-order conditions directly, without introducing additional parameters to be estimated. This enables researchers to use variation in tax rates to distinguish models of conduct even when other instruments fail, or to falsify classes of models that other instruments may not. In addition, as shown in Propositions 1 and 2, the marginal effect of an instrument on prices plays an important role in its ability to distinguish between models of conduct, and there is a long line of empirical literature documenting the strong effects of tax rate changes on prices (see, e.g., Butters, Sacks, and Seo (2022) and the references therein).<sup>20</sup>

## 5 Instrument Selection for Counterfactuals

Researchers often want to know the correct model of conduct not for its own sake, but to perform counterfactual analyses. Different counterfactual exercises depend on different features of the model; our falsification framework can identify instruments that target these particular features, making the counterfactuals more credible. As an illustration, we consider the problem of maximizing tax revenue by learning the Laffer curve, which has garnered recent interest (Miravete et al., 2018; Hollenbeck and Uetake, 2021; Brugués and De Simone, 2024).<sup>21</sup>

<sup>20</sup>In contrast, DellaVigna and Gentzkow (2019) provide evidence that some types of demand shifters may only have weak effects on prices in the data-generating process. This also appears to be an issue for other common conduct instruments, leading to inferential problems (Backus et al., 2021; Duarte et al., 2024).

<sup>21</sup>The insights in this section readily extend to other optimal tax policy settings: for example, if the government wished to maximize a convex combination of revenue and consumer surplus.

## 5.1 Learning the Laffer Curve

Counterfactual predictions are essential to understanding the Laffer curve, as taxes usually have limited support in the data. The policymaker must use the model to forecast revenue at unobserved tax rates. We limit attention to the optimal tax in a single market,  $t$ , for expositional parsimony. Government revenues in market  $t$  for any tax rates  $(\nu, \tau)$  are

$$R_t(\nu, \tau) = \tau \sum_j s_{jt}(p_{0t}(\nu, \tau)) + (1 - \nu) \sum_j p_{0jt}(\nu, \tau) s_{jt}(p_{0t}(\nu, \tau)). \quad (5)$$

This expression emphasizes that market shares depend on prices, and prices on tax rates; all other primitives are held constant and suppressed. The observed price  $p_t = p_{0t}(\nu_t, \tau_t)$  corresponds to the current tax rates  $(\nu_t, \tau_t)$ . The policymaker knows the revenue at observed tax rates, but must use a model  $m$  to predict counterfactual prices  $p_{mt}(\nu, \tau)$  and therefore revenue at other rates. Recent literature (Miravete et al., 2018; Hollenbeck and Uetake, 2021) shows that the Laffer curve can differ substantially across different models. The policymaker, not observing the true conduct model, wants the model that best matches the true Laffer curve. Our framework sheds light on the link between falsification by particular instruments and the counterfactual Laffer curve.

We assume the government optimizes one tax rate, keeping the other fixed. Our application to the Washington cannabis market exemplifies this: the government only uses an ad valorem tax. We analyze the Laffer curves for unit and ad valorem tax rates separately.

**Unit Tax:** Fix the ad valorem tax at its observed level  $\nu = \nu_t$  and allow the unit tax to vary. Equation (5) reveals which aspect of firm conduct determines the Laffer curve. Assuming the demand function  $s_{jt}(\cdot)$  is identified, we need only identify  $p_{0t}(\nu_t, \cdot)$ , the effect of unit tax on prices. Applying the Implicit Function Theorem to Equation (3) yields  $\frac{dp_{0t}}{d\tau} = \frac{1}{\nu_t} P_{0t}(\tau) \iota$ , where  $\iota$  is a  $J$ -vector of ones.<sup>22</sup>

**Corollary 5.** *Under Assumptions 1-6 and 9, if  $P_{mt}\iota = P_{0t}\iota$  everywhere, then model  $m$  produces the true Laffer curve for the unit tax.*

The proof, detailed in the Appendix A, can be summarized as follows. All models are constructed to rationalize observed prices, so  $p_{mt}(\tau_t) = p_{0t}(\tau_t) = p_t$ . With  $\nu_t$  fixed,  $P_{mt}\iota = P_{0t}\iota$

---

<sup>22</sup>Formally, stack Equation (3) and rewrite it as  $\nu_t p_t - \nu_t \Delta_{0t} - \tau_t \iota = 0$ . With  $\nu_t$  fixed, define  $F(p_t, \tau_t)$  as the left-hand side, so  $p_{0t}(\tau)$  is implicitly defined as the solution to  $F(p_{0t}(\tau), \tau) = 0$ . The Implicit Function Theorem then gives  $\frac{dp_{0t}}{d\tau} = -\left(\frac{\partial F}{\partial p_t}\right)^{-1} \frac{\partial F}{\partial \tau_t} = -\left(\nu_t I - \nu_t \frac{\partial \Delta_{0t}}{\partial p_t}\right)^{-1} (-\iota)$  under Assumption 9. As shown in the text,  $P_{0t} = \left(I - \frac{\partial \Delta_{0t}}{\partial p_t}\right)^{-1}$ , and therefore  $\frac{dp_{0t}}{d\tau} = \frac{1}{\nu_t} P_{0t} \iota$ .

is sufficient for the derivatives of  $p_{mt}(\cdot)$  and  $p_{0t}(\cdot)$  to match everywhere, making these functions identical and resulting in identical Laffer curves. Thus, Corollary 5 shows that model  $m$  yields the correct optimal tax when the row sums of the model's pass-through matrix equals that of the data-generating pass-through matrix, a weaker requirement than  $P_{mt} = P_{0t}$ .

Using the unit tax as an instrument creates a close link between the falsification condition in Proposition 2 and the Laffer curve equivalence condition in Corollary 5. With a single unit tax applying to all products, the falsification condition in Equation (4) simplifies to:  $E [P_{mt}^{-1} (P_{0t\iota} - P_{mt\iota}) \mid w_t, \tau_t] \neq 0$ . Therefore, model  $m$  is not falsified when  $P_{mt\iota} = P_{0t\iota}$  almost surely on the data's support. Thus, using the unit tax as the instrument directly targets the most relevant part of the pass-through matrix for the unit tax policy counterfactual. Example 7 further illustrates this connection.

*Example 7:* Consider the simple environment of Example 1 where the true model is Keystone pricing and the alternative model is Bertrand. Under Proposition 2, variation in a unit tax, which moves own and rival cost, falsifies Bertrand as  $P_{Bt\iota} < 1 < 2 = P_{Kt\iota}$ . Further, since  $P_{Bt\iota} \neq P_{Kt\iota}$ , the Laffer curve under Bertrand differs from the truth. Now consider falsifying a shifted Keystone model ( $m = S$ ) whereby  $p_{jt} = 2c_{jt} + \zeta_{jt}$ , where  $\zeta_{jt}$  is a constant. Notice that, since  $P_{St} = P_{Kt}$  and therefore  $P_{St\iota} = P_{Kt\iota}$ , the shifted Keystone model is not falsified by the unit tax instrument. However, because the pass-through matrices are the same, the two models imply the same unit tax Laffer curve. Thus, while variation in the unit tax may not falsify all wrong models, it will permit the researcher to learn the true Laffer curve. •

**Ad Valorem Tax:** Now fixing the unit tax at  $\tau = \tau_t$ , we examine the Laffer curve for the ad valorem tax. With demand identified, accurately capturing the tax's effect on prices is once again the key step. Applying the Implicit Function Theorem to Equation (3) yields  $\frac{dp_{0t}}{d\nu} = \frac{1}{\nu} P_{0t}(\nu)(\Delta_{0t}(\nu) - p_{0t}(\nu))$ , where only  $\nu$  varies and we suppress dependence on other arguments. The key term to determine the Laffer curve is  $P_{0t}(\Delta_{0t} - p_{0t})$ . This resembles the unit tax case, with  $\Delta_{0t} - p_{0t}$  replacing  $\iota$  due to  $\nu_t$ 's multiplicative entry in the equilibrium condition. As before, imposing model  $m$  will result in a potentially different implied Laffer curve.

**Corollary 6.** *Under Assumptions 1-6 and 9, if  $P_{mt}(\Delta_{mt} - p_{mt}) = P_{0t}(\Delta_{0t} - p_{0t})$  whenever  $p_{mt} = p_{0t}$ , then model  $m$  produces the true Laffer curve for the ad valorem tax.*

A consequence of Corollary 6 is that any rule-of-thumb model where markups are a constant fraction of prices – such as Keystone and marginal cost pricing – will yield the same Laffer curve.<sup>23</sup> For model  $m$  in the rule-of-thumb class,  $\Delta_{mt} = \psi_m p_t$ , and  $P_{mt} = \frac{1}{1-\psi_m} I$  and  $(\Delta_{mt} - p_t) = (\psi_m - 1)p_t$  so that  $P_{mt}(\Delta_{mt} - p_t) = -p_t$ , which does not depend on  $\psi_m$ .

<sup>23</sup>For Keystone,  $\Delta_{Kt} = 0.5p_t$ , while for marginal cost pricing,  $\Delta_{MCt} = 0 \times p_t$ .

Using the ad valorem tax as a conduct instrument links falsification to the optimal tax counterfactual. Because  $p_t = p_{0t} = p_{mt}$  in the data, the falsification condition in Equation (4) simplifies to  $E \left[ P_{mt}^{-1} \left( P_{0t} (\Delta_{0t} - p_t) - P_{mt} (\Delta_{mt} - p_t) \right) \mid w_t, \nu_t \right] \neq 0$ . If  $P_{mt} (\Delta_{mt} - p_t) = P_{0t} (\Delta_{0t} - p_t)$  almost surely in the support of the data, model  $m$  is not falsified. Thus, the ad valorem tax instrument targets the most relevant objects for the tax counterfactual.

## 5.2 Counterfactual-Relevant Instruments: A Discussion

Corollaries 5 and 6 show that the tax rate itself is the most relevant instrument to learn conduct to find the Laffer curve, as it targets the most crucial aspects of pass-through and markups. However, this raises a question: if we observe both tax rates and market outcomes, why use a structural model instead of simply identifying the tax rate with the highest observed revenue? In real-world data, variation in observed tax rates is often limited, while finding the optimal tax requires inferring revenue for every potential tax rate. Using the existing variation in taxes allows us to learn the model of conduct that best fits the features of the data relevant to the Laffer curve. However, when the data does not contain any variation in the tax rate of interest, a researcher can use our framework to understand the tradeoffs involved in selecting another instrument (i.e., the other tax rate, rival cost shifters, or product characteristics) to learn conduct and subsequently the Laffer curve.

To illustrate how our framework informs selecting “second best” instruments, consider the case of setting an optimal unit tax. The key object is  $P_{0t}\iota$ , which the unit tax instrument directly targets. Insofar as variation in the ad valorem tax rate exists, it may seem intuitive to use variation in the “other tax” to learn conduct and the Laffer curve. However, our results above show that the ad valorem tax targets  $P_{0t} (\Delta_{0t} - p_{0t})$ , making it generally difficult to target  $P_{0t}\iota$  with these instruments. In the special case where marginal costs are the same for all products, so that  $c_{0jt} = c_{0t}^*$  and  $c_{mjt} = c_{mt}^*$  for all  $(j, t)$ , falsifiability requires  $E \left[ \frac{1}{\nu_t^2} P_{mt}^{-1} (c_{0t}^* P_{0t}\iota - c_{mt}^* P_{mt}\iota) \mid w_t, \nu_t \right] \neq 0$ . At best, the ad valorem tax instrument can only target  $P_{0t}\iota$  up to an unknown, market-varying scale factor  $c_{0t}^*$ . Thus, it appears that the ad valorem tax instrument cannot be first-best for the unit tax Laffer curve.

With demand side and cost side instruments, the situation is more optimistic, though there are tradeoffs. While our results in Section 4 show that demand and cost side instruments do not generally target  $P_{0t}\iota$ , cases exist where both target  $P_{0t}$  directly. Outside of the knife-edge case where the indirect and direct effects perfectly offset, demand side instruments constructed with own and rival product characteristics simultaneously target both  $P_{0t}$  and the direct effect. Furthermore, when the demand system satisfies Assumption 8 and the conduct model satisfies Assumption 7, the demand side instruments target only  $P_{0t}$ . For

cost side instruments, consider the special case where marginal cost depends additively a special shifter  $w_{jt}^*$  such that  $c_{jt} = \bar{c}_{jt} + \gamma w_{jt}^* + \omega_{0jt}$ . When a researcher both observes  $w^*$  and knows the value of  $\gamma$  (e.g., for observed wholesale prices,  $\gamma = 1$ ), they can use variation in own and rival  $w^*$  to construct instruments. These instruments function like product-specific unit taxes, and it is straightforward to show that they target  $P_{0t}$ .

The benefit of targeting  $P_{0t}$  is that models where  $P_{mt} = P_{0t}$  will deliver the true Laffer curve. The downside is that targeting the full pass-through matrix  $P_{0t}$  is stricter than targeting  $P_{0t\ell}$ . This is potentially problematic because it implies that such instruments may falsify models which deliver the true Laffer curve. One can circumvent this issue by constructing an instrument from the market-level average of  $x_{jt}$  or  $w_{jt}^*$  (across own and rival products). Under Assumptions 7 and 8, using the average value of product characteristics  $\bar{x}_t$  as an instrument yields the falsification condition  $\frac{\beta}{\alpha} E \left[ P_{mt}^{-1} (P_{0t\ell} - P_{mt\ell}) \mid w_t, \tilde{x}_t, \bar{x}_t \right] \neq 0$ , where  $\tilde{x}_{jt} = x_{jt} - \bar{x}_t$ . Similarly, when the special shifter  $w_{jt}^*$  is observed and  $\gamma$  is known, using  $\bar{w}_t^*$  as the instrument directly targets  $P_{0t\ell}$  as the falsification condition becomes  $\gamma E \left[ P_{mt}^{-1} (P_{0t\ell} - P_{mt\ell}) \mid \tilde{w}_t, \bar{w}_t \right] \neq 0$ . In either case, the market-level average instrument targets  $P_{0t\ell}$  directly.

Consider now the case of setting an optimal ad valorem tax. Both pass-through and the level of markups determine the ad valorem Laffer curve, and no instrument other than ad valorem taxes targets both of these objects. Which instrument is second-best in this case depends on the relative importance of indirect and level effects, which will vary by application.

Overall, these examples show how researchers can use our falsification framework to understand the tradeoffs inherent to various instruments. In the case of Laffer curves, imposing additional assumptions – e.g., linearity, known coefficients, and index restrictions – can enable researchers to effectively replicate the first-best conduct instrument. These assumptions can be supported by economic intuition and institutional knowledge, and in some cases they are testable.<sup>24</sup> Whether such a replication is possible depends on the counterfactual at hand.

## 6 Monte Carlo Simulations

We further illustrate the main results of the paper with simulations tied to the examples described in the preceding sections. To align the environment with the falsification framework, we simulate data for 50,000 markets using the simulation class in PyBLP (Conlon and Gortmaker, 2020). In each market  $t$ , the number of single-product firms  $J_t$  is a randomly chosen integer between two to ten, leaving us with 319,719 observations in the sample. We adopt a simple logit demand system, in line with our falsification examples.

The parameterization of demand results in a mean own price elasticity of -8.59 and

---

<sup>24</sup>For example, the demand side index restriction in Assumption 8 is testable.

diversion to the outside option of 0.37. On the supply side, we assume that the marginal costs are a linear function of two observed cost shifters, which are excluded from demand, and of  $\omega_{0jt}$ , the true unobserved cost shock. The government levies both a unit tax ( $\tau_t$ ) and an ad valorem tax ( $v_t$ ) on all products in market  $t$ . The unit tax is remitted by the firms while consumers remit the ad valorem tax. We assume that the true model of conduct is Keystone pricing, whereby firms set tax exclusive prices as twice their marginal cost, or  $\nu_t p_{jt} = 2c_{jt} + 2\tau_t$ . Full details of our simulation environment are available in Appendix D.

## 6.1 Testing with Rivers and Vuong (2002)

The researcher considers a menu of four models: the true Keystone model ( $m = K$ ), Bertrand ( $m = B$ ), Cournot ( $m = C$ ), and a shifted Keystone model ( $m = S$ ). For the shifted Keystone model,  $\nu_t p_{jt} = \nu_t \eta + 2c_{mjt} + 2\tau_t$ , where a fixed  $\eta = 1$  is added to markups for all products in all markets. To make inferences on firm conduct in finite samples, the researcher adopts the model selection test in Rivers and Vuong (2002) (RV). For the RV test, the null hypothesis states that two competing models of conduct  $m = 1, 2$  have the same lack-of-fit, or  $H_0^{\text{RV}} : Q_1 = Q_2$ , while the alternatives correspond to superior fit of one of the two models,  $H_1^{\text{RV}} : Q_1 < Q_2$  and  $H_2^{\text{RV}} : Q_2 < Q_1$ . With this formulation of the hypotheses, the testing procedure determines which of the two models has the smallest lack of fit.

As in Duarte et al. (2024), we express  $Q_m$  as a GMM objective function  $Q_m = g_m' W g_m$ , where  $g_m = E[z_{jt} \omega_{mjt}]$  and  $W = E[z_{jt} z_{jt}']^{-1}$  is the weight matrix. For this GMM measure of fit, the RV test statistic is then

$$T^{\text{RV}} = \frac{\sqrt{n}(\hat{Q}_1 - \hat{Q}_2)}{\hat{\sigma}_{\text{RV}}},$$

where  $\hat{Q}_m = \hat{g}_m' \hat{W} \hat{g}_m$ ,  $\hat{g}_m = n^{-1} \sum z_{jt}' \hat{\omega}_{mjt}$ ,  $\hat{W} = n(\hat{z}' \hat{z})^{-1}$ , and  $\hat{\sigma}_{\text{RV}}^2$  is an estimator for the asymptotic variance of  $\sqrt{n}(\hat{Q}_1 - \hat{Q}_2)$ .<sup>25</sup>

By imposing the true demand system, and considering the true model as one of the alternatives, the RV test has a close connection to the falsification framework. Suppose the true model is model 1; then  $\hat{Q}_1 \simeq 0$  in large samples. Then, the numerator of the RV test coincides with the falsifiable restriction for model 2, and rejection of the RV null in favor of model 1 coincides with the falsification of model 2. However, when both models are not falsified, the RV test statistic is degenerate, causing severe inferential problems. Thus, the  $F$ -statistic diagnostic from Duarte et al. (2024) can be used as finite-sample evidence of instrument relevance. In addition to this ex-post diagnostic, the results in Sections 4 and

---

<sup>25</sup>See Duarte et al. (2024) for a variance estimator accounting for randomness in  $\hat{W}$  and  $\hat{g}_m$ .

5 show how a researcher can use the economics of pass-through implied by the models of conduct being tested to choose ex-ante potentially relevant instruments.

To perform testing, we consider four sets of instruments: the unit tax  $\tau_t$ , the ad valorem tax  $v_t$ , cost side instruments (the sum of rival cost shifters in a market, denoted as “Riv. Cost”), and demand side instruments (the number of rival products in a market and the sum of rival product characteristics, denoted as “Riv. BLP”), which we also pair with own characteristics (“Riv. BLP + Own  $x$ ’s”). We run the RV test separately with each instrument set.

## 6.2 Simulation Results

**Test Results:** Table 1 presents the results of testing all pairs of models. Examples 1-6 suggest that when testing Bertrand against Keystone, all instruments should be relevant.<sup>26</sup>

TABLE 1: Test Results for Simulated Data

| Statistic                                     | Instruments:    |                       |                  |                 |                              |
|-----------------------------------------------|-----------------|-----------------------|------------------|-----------------|------------------------------|
|                                               | Unit Tax<br>(1) | Ad Valorem Tax<br>(2) | Riv. Cost<br>(3) | Riv. BLP<br>(4) | Riv. BLP + Own $x$ ’s<br>(5) |
| <b>Panel A: Bertrand vs. Keystone</b>         |                 |                       |                  |                 |                              |
| $T^{RV}$                                      | 25.2            | 16.1                  | 9.8              | 16.3            | 29.4                         |
|                                               | ***             | ***                   | ***              | ***             | ***                          |
| $F$                                           | 249.3           | 138.1                 | 21.2             | 44.8            | 157.6                        |
|                                               | ††† ^^^         | ††† ^^^               | ††† ^^^          | ††† ^^^         | ††† ^^^                      |
| <b>Panel B: Cournot vs. Keystone</b>          |                 |                       |                  |                 |                              |
| $T^{RV}$                                      | 23.7            | 23.4                  | −0.0             | 0.7             | 40.8                         |
|                                               | ***             | ***                   |                  |                 | ***                          |
| $F$                                           | 47.8            | 46.8                  | 0.1              | 0.0             | 66.8                         |
|                                               | ††† ^^^         | ††† ^^^               | †††              | †††             | ††† ^^^                      |
| <b>Panel C: Shifted Keystone vs. Keystone</b> |                 |                       |                  |                 |                              |
| $T^{RV}$                                      | −0.1            | 4.8                   | 0.6              | 0.5             | 0.5                          |
|                                               |                 | ***                   |                  |                 |                              |

The table reports, for each set of instruments and pair of models, the RV test statistics  $T^{RV}$ . Panels A and B also report the effective  $F$ -statistic (Duarte et al., 2024); we don’t report it in Panel C as the pair of models does not satisfy Assumption 2 in Duarte et al. (2024). Columns 1-5 correspond to different instruments; panels A-C correspond to three different pairs of models. A positive RV test statistic suggests a better fit of Keystone. The symbol \*\*\* indicates rejection of the null of equal fit 0.01 confidence level. The symbols ††† and ^^^ indicated that  $F$  is above the appropriate critical values for worst-case size below 0.075, and best-case power above 0.95, respectively. Both  $T^{RV}$  and the  $F$ -statistics account for market level clustering.  $n = 319,719$ .

We see in Panel A that the RV test statistics are all positive for this pair of models and are well above 1.96, so that the null is rejected in favor of Keystone. When testing Bertrand

<sup>26</sup> A table in Appendix E summarizes instrument relevance results for these examples.

against Keystone, the  $F$ -statistics are all larger than the critical values for worst-case size of 0.075 and best-case power of 0.95, suggesting all instruments are relevant.

By contrast, Example 3 suggests that variation in rival product characteristics and rival cost shifters are irrelevant for falsifying Cournot against Keystone when demand is logit. We see in Columns 3-4 of Panel B that these instruments have virtually no power and the null is not rejected. Example 4 however suggests that variation in own product characteristics, which can pick up on differences in the main diagonal of the pass-through matrices, should falsify Cournot. In Column 5, we see that including own product characteristics creates relevant instruments which reject the null in favor of the true Keystone model. Examples 5 and 6 suggest that the tax instruments are relevant: Columns 1-2 show that these instruments are strong for size and power and reject the null in favor of Keystone.

Finally, when testing the shifted Keystone model against Keystone (Panel C), both models have the same pass-through matrices as discussed in Example 7. Thus, we expect the unit tax, product characteristics (own and rival), and rival cost shifters (Columns 1,3-5) to be irrelevant, and  $T^{RV}$  is close to zero in those cases. However, the level of the markups differs across the two models. Ad valorem tax instruments (Column 2), which can falsify a wrong model based on the level effect, strongly reject the shifted Keystone model in favor of Keystone.

**Laffer Curve:** Next, we consider the tax implications of the models tested above. To mimic the problem facing a state government setting a single tax rate across many markets, we randomly select 500 of the 50,000 markets in our simulated data. In Panel A of Figure 1, we compute the Laffer curve for the unit tax, holding the ad valorem tax fixed at the levels in the simulated data. To obtain the Laffer curve under a given model of conduct, we first back out  $c_{mjt}$ , the implied marginal cost under that model. Holding demand and cost fixed, we then solve for equilibrium at each level of the tax and compute the associated revenue.

From the results in Table 1, we find that Keystone and shifted Keystone are the two models that are not falsified by unit tax instruments, as they feature the same pass-through matrices. Furthermore, in line with Corollary 5, we see in Panel A of Figure 1 that these models imply the same Laffer curve. Conversely, since Bertrand and Cournot yield different Laffer curves than the true one, they have different  $P_{mt}$  than Keystone, which permits the falsification of these models with unit tax instruments (Panels A and B of Table 1).

In Panel B of Figure 1, we compute the Laffer curve for the ad valorem tax, holding the unit tax fixed at their simulated levels and following the same steps outlined for Panel A. In Table 1, we find that Keystone is the only model not falsified by ad valorem tax instruments. This is in line with Corollary 6: the other models all produce a different Laffer curve, thus have different  $P_{mt}(\Delta_{mt} - p_t)$ , and are therefore falsified by ad valorem tax instruments.

These results illustrate three ways in which our falsification framework can be used. First,

FIGURE 1: Laffer Curve for Unit and Ad Valorem Taxes

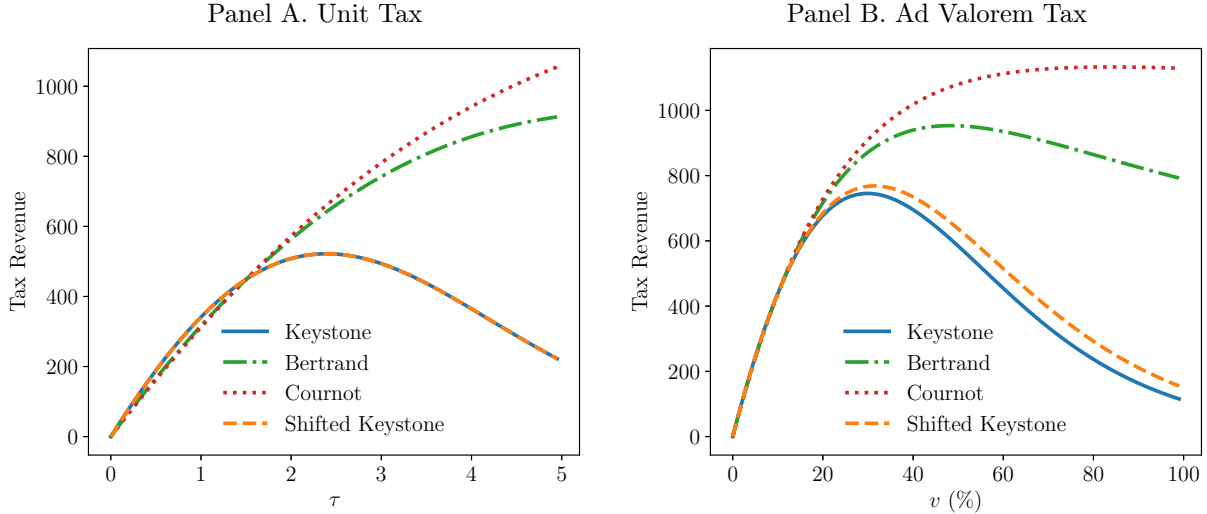


Figure illustrates the Laffer curve implied by different models of conduct. In Panel A, we hold fixed the ad valorem tax in each market and vary the unit tax from 0 to 5. In Panel B, we hold fixed the unit tax but vary the ad valorem tax from zero to one. To compute the curves in each panel, we simulate counterfactual prices, quantities, and government revenue at different levels of the appropriate tax for each candidate model.

the falsification results in Section 4 enable ex-ante instrument selection to avoid irrelevant instruments. For instance, when testing Keystone versus Cournot, our results suggested that cost side instruments would be irrelevant. Second, our framework can motivate instrument choice for specific models. For instance, we found ad valorem tax instruments to be particularly useful to falsify the shifted Keystone model, which is otherwise hard to distinguish from Keystone. Finally, when a researcher targets a specific counterfactual, as in Section 5, they can use our framework to tailor the instrument choice for learning the optimal policy.

With these takeaways in mind, we now turn to an empirical setting where conduct is unknown, and policymakers are interested in setting revenue-maximizing taxes.

## 7 Empirical Application

As an application of our results, we consider the problem of setting the optimal ad valorem tax in Washington State’s marijuana market, which is well-suited for several reasons. First, setting the optimal tax is an important consideration in this market, as revenues are substantial. Second, optimal taxes depend on conduct (Miravete et al., 2018; Hollenbeck and Uetake, 2021), and it is unclear how well retailers’ conduct in this market aligns with standard structural models (Escudero, 2018; Hollenbeck et al., 2024). Third, our results suggest that for a tax exercise, conduct would best be learned using an ad valorem tax instrument, which is available in this market through local cannabis tax variation.

## 7.1 Institutional Background and Data

Washington legalized cannabis for medical use in 1998, and in 2012, through voter approval and passage of I-502, was among the first states to legalize recreational cannabis.<sup>27</sup> In July of 2014, retail sales of cannabis began in Washington.

The state regulates the cannabis industry in several ways. First, there is a cap on the number of retailers.<sup>28</sup> Second, the state restricts vertical relations between producers (growers), processors, and retailers. While producers and processors — those who convert marijuana plants into various products — can integrate with one another, neither can integrate with a retailer. Only retailers can sell cannabis products directly to consumers. Third, the state assesses an ad valorem tax on cannabis sales. When cannabis was first legalized, the state levied a 25% tax at each point of sale — from producers to processors, processors to retailers, and retailers to consumers. In July 2015, this changed to a 37% tax on the final sale to consumers only.<sup>29</sup> Additionally, local municipalities imposed their own ad valorem taxes, which range from seven to ten percent.

In part to enforce tax collection, the state tracks detailed “seed-to-sale” data on all cannabis products. Our empirical application benefits from these unusually detailed data. We construct a dataset containing product shares and prices for many markets. The underlying data come from BioTrack, an administrative data collection software that tracks cannabis from plant production and processing through retail (“seed-to-sale”).<sup>30</sup> These data include the universe of wholesaler and retailer transactions for cannabis products. Thus, for every transaction, we observe prices paid both by consumers to retailers as well as by retailers to wholesalers. Biotrack also collects supplemental information on the organizations in the market, from which we observe when retailers began operating, and lab testing of plants, from which we observe cannabis product CBD and THC content.

Due to the difficulty of comparing package size weights between different product types (e.g. liquid, extracts, and solids), we only keep products labeled as “usable” marijuana (dried leaves and flowers), which represent 72% of all transactions in the data. Within this category, we subset to package sizes of 1 and 3.5 grams (following [Escudero, 2018](#)), which are by far the most popular, accounting for 60% of usable product revenue.<sup>31</sup> Features of the cannabis

---

<sup>27</sup>As of 2024, 38 states have legalized cannabis for medical use and 24 for recreational use.

<sup>28</sup>The Washington State Liquor and Cannabis Board initially limited the number of retailer licenses to 334, and allocated licenses across counties proportional to population. In jurisdictions with too many applications, licenses were assigned via lottery. In January 2016, the state increased the number of retail licenses to 556. The state also caps the number of each licensee’s retail locations at three.

<sup>29</sup>[Hansen, Miller, and Weber \(2022\)](#) explore the effects of this tax change on the tendency for producers and processors to vertically integrate.

<sup>30</sup>These tracking efforts are, in part, to thwart sales in neighboring states where cannabis was illegal.

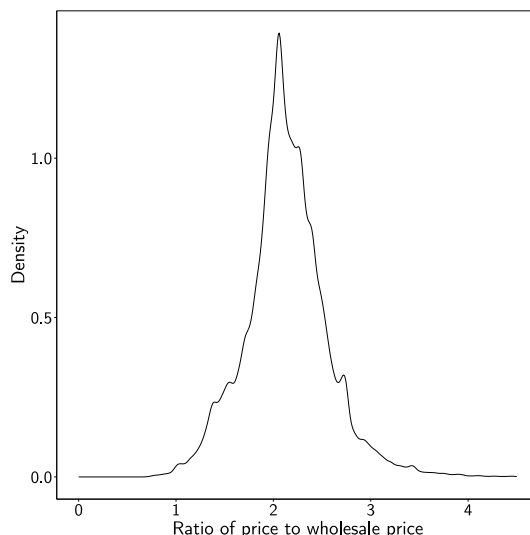
<sup>31</sup>See Table 4 in Appendix F for a summary of all data cleaning steps.

market complicate the observed prices in the raw data. Because cannabis is illegal at the federal level, retailers accept cash payments and manually record transactions, resulting in inconsistencies regarding whether reported prices included the tax. [Hollenbeck and Uetake \(2021\)](#) correct for this, and we use their reconstructed tax-inclusive retail prices.<sup>32</sup>

We define a product to be a package size from a processor sold at a retailer in a month and markets to be city-month pairs. We calculate a product’s price as the sales volume-weighted average price. We define the market size to be twice the maximum amount ever purchased in a given geographic market. We further draw 2000 consumers’ incomes from the empirical distribution of income in each city, which we obtain from the US Census. The data include prices and shares for 1,428 products across 2,639 markets from August 2014 to May 2017.

Appendix Table 5 displays summary statistics of the data used in estimation. Wholesalers charge an average of \$3.60 for a gram of cannabis, while retailers charge \$7.60, on average. While the average market has over 100 cannabis processors, retail market quotas restrict the average market to 14 retailers. We follow the literature ([Escudero, 2018](#); [Hollenbeck and Uetake, 2021](#)) in concluding from these facts that wholesaler market power is unlikely to be important, and we focus our attention on how retailers set prices instead.

FIGURE 2: Ratio of Retail and Wholesale Price



The figure plots densities of the ratio of retail and wholesale prices for each product in our data.

Figure 2 graphs the ratio of retail unit prices to wholesale unit prices. We see a mass of ratios concentrated around two, lending credence to the hypothesis that firms engage in Keystone pricing. A formal test follows in Section 7.3.

---

<sup>32</sup>Additional details on data construction can be found in Appendix A of [Hollenbeck and Uetake \(2021\)](#).

## 7.2 Demand Estimation

We follow [Hollenbeck and Uetake \(2021\)](#) in adopting a random coefficient nested logit demand model ([Grigolon and Verboven, 2014](#)). We briefly summarize our demand model here, and provide a more detailed exposition in Appendix F.

Product characteristics include a constant, package size, THC and CBD levels (and their values squared), and the log of the number of products offered in the store to capture variation in shelf space across stores. We allow consumer preferences for characteristics and prices to vary with individual-level income. We also include fixed effects at the retailer, processor, and year-month levels. Our nesting structure includes all inside goods in one nest.

Following [Hollenbeck and Uetake \(2021\)](#) we construct demand instruments from several sources of exogeneous variation. These include the number of products sold at competing dispensaries in the market, also interacted with the mean income, BLP-style instruments (sum of THC and CBD for products within a store, and for products in all other stores), and exogenous own cost shifters (rainfall and temperature in the region of production and their lags). We estimate demand using PyBLP ([Conlon and Gortmaker, 2020](#)).

Results for demand estimation are reported in Appendix Table 6. Similar to [Hollenbeck and Uetake \(2021\)](#), we find a positive interaction between income and price and a negative interaction between income and the constant. We also find that consumers prefer products with a higher level of THC, though at a declining rate, while they dislike products with higher levels of CBD. The median own-price elasticity in our sample is -6.451 suggesting a fair degree of substitution across strands of usable marijuana. Inspection of the aggregate elasticity allows comparison to the results in [Hollenbeck and Uetake \(2021\)](#), who aggregate products of the same type within a store, and to demand for related products (e.g., liquor). Our aggregate elasticity of -3.32 (-2.9 when weighted by market size) is in line with median own-price elasticity in [Hollenbeck and Uetake \(2021\)](#) (-2.89) and broadly similar to the aggregate elasticity in the Pennsylvania liquor market found in [Miravete, Seim, and Thurk \(2020\)](#) (-2.48).

One potential concern is that our implementation of the RCNL model restricts the range of curvature of demand ([Miravete et al., 2024](#)). In turn, misspecifying demand in this way can lead to restrictions on the pass-through implied by certain models of conduct (e.g., Bertrand). Because we show, in the previous section, that differences in pass-through matrices help distinguish different models of conduct, this potential misspecification could affect our test results. [Miravete et al. \(2024\)](#) allow for greater flexibility in modeling heterogeneous price sensitivity by incorporating a Box-Cox transformation of the income interaction. They show that this can greatly increase the range of curvature implied by discrete choice demand models. Without micro-data, we lack sufficient identifying power to pin down the additional

parameters required by this approach. Instead, in Appendix G, we calibrate different approximations for the Box-Cox transformations for income and include them in the RCNL demand system. We find that our testing results are unaffected.

### 7.3 Test for Conduct

**Models of Conduct:** We consider three models of conduct that are discussed in previous literature for this market: (i) *Marginal Cost Pricing*—retailers choose tax-exclusive prices equal to marginal cost; (ii) *Keystone Pricing*—retailers choose tax-exclusive prices equal to twice marginal cost; and (iii) *Bertrand Pricing (Nash Price Setting)*—retailers set prices competitively to maximize profits in the complete information pricing game (considered in Hollenbeck and Uetake, 2021). Escudero (2018) provides some descriptive evidence that Keystone pricing is plausible in the market during our sample period, while Hollenbeck et al. (2024) provide evidence of strategic interactions in a later period. It is thus unclear *a priori* which model best fits the data during our sample period.

For the first two models, implied costs and markups can be immediately computed from observed prices. Given our demand estimates, we can also compute  $\Delta_B$  and  $c_B$  for the Bertrand model. We specify marginal cost as the sum of observed wholesale prices  $p^W$  and additional retail costs. In turn, these include a linear function of observed shifters  $w$  and an unobserved shock. The vector  $w_{jt}$  includes product fixed effects.<sup>33</sup>

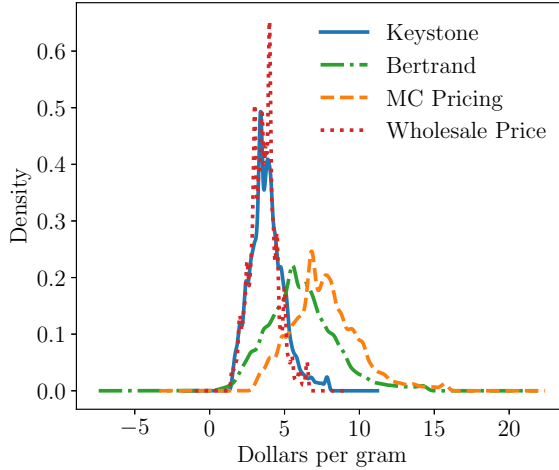
**Inspection of Implied Costs and Markups:** Before performing formal testing, we inspect the data and the implications of different models for markups and costs. Given that we observe wholesale prices, and that these likely make up a large fraction of the marginal cost of selling a product, it is natural to compare the model-implied costs to these prices. Figure 3 reports the distributions of implied costs for all models, as well as the observed wholesale prices in dollars per gram. We notice first that Keystone implies lower marginal costs than Bertrand which in turn implies lower marginal costs than marginal cost pricing. While the distribution of the Keystone implied costs is close to the distribution of wholesale prices, the other models' implied cost distributions are far from wholesale prices.

The descriptive statistics reported in Table 2 further shed light on the differences across the three models. The fit as measured by RMSE of implied costs with respect to observed wholesale prices is substantially better for Keystone. Moreover, wholesale prices represent on average 98% of the Keystone total marginal cost. In contrast, marginal cost pricing and Bertrand imply that wholesale prices represents only 49% and 65% of retail marginal costs, respectively. Inspection of markups reveals that while firms have market power under Bertrand,

---

<sup>33</sup>In Appendix G, we show the robustness of our test results to alternative specifications of cost.

FIGURE 3: Model Implied Marginal Costs



Distributions of marginal costs implied by different models of conduct along with observed wholesale prices

markups are even higher under Keystone pricing. While not conclusive, our descriptive analysis suggests that the Keystone model is the most reasonable of the three and that the model of marginal cost pricing is implausible in this market. Additionally, our findings in Section 5.1 indicate that both Keystone and marginal cost pricing will produce the same counterfactual Laffer curve for the ad valorem tax because they have identical ad valorem tax pass-through. Thus, going forward, we focus on distinguishing Bertrand from Keystone.

TABLE 2: Descriptive Statistics of Model Implied Markups and Costs

| Statistic                       | MC Pricing | Keystone | Bertrand |
|---------------------------------|------------|----------|----------|
| RMSE of $c_m$ and $p^W$         | 4.38       | 0.78     | 3.04     |
| Mean $p^W/c_m$                  | 0.49       | 0.98     | 0.65     |
| Mean $\Delta_m/p$               | 0          | 0.50     | 0.22     |
| Mean $\Delta_m$ (\$)            | 0          | 3.84     | 1.54     |
| 25th percentile $\Delta_m$ (\$) | 0          | 3.10     | 1.20     |
| 50th percentile $\Delta_m$ (\$) | 0          | 3.75     | 1.30     |
| 75th percentile $\Delta_m$ (\$) | 0          | 4.45     | 1.69     |

We report descriptive statistics of measures of model implied markups and costs. Each column corresponds to a different model of conduct.

**Implementation of the Test:** To make inference on firm conduct in finite samples, we adopt the model selection test in Rivers and Vuong (2002) (RV) introduced in Section 6, which requires instruments. Because we have variation in the data in state-wide and local ad valorem taxes, our results in Section 5 suggest that for an optimal ad valorem taxation counterfactual, we should use ad valorem tax instruments. However, if the researcher has a different objective, such as learning conduct for its own sake, they may want to consider

different sets of relevant instruments. In addition to taxes, three other sources of variation in the data should be relevant for testing. These are (i) rival cost shifters - RC (sum of rainfall, temperature, and their lags); (ii) product characteristics - PC (own and rival number of products, total own and rival CBD and THC); (iii) wholesale prices - WP. In our setting, wholesale prices are functionally equivalent to product-specific unit tax rates.

While all of these instruments are exogenous under standard assumptions (including marginal cost pricing upstream), their relevance can be analyzed in light of our falsification framework. Assuming that the true model is one of the two candidates we test in this application, any of these instruments is plausibly relevant, i.e. capable of falsifying the wrong model, as shown in the previous examples. However, we also learn that different instruments leverage different features of the pass-through matrix to distinguish models of conduct. Thus, using each set of instruments separately to perform the test will provide robust evidence.

Beyond our falsification results, the ability of different sets of instruments to falsify conduct in a finite sample will depend on the variation present in the data. To assess the strength of instruments in the data, we report values of the  $F$ -statistic for the RV test developed in [Duarte et al. \(2024\)](#).

**Test Results:** We test the Bertrand and Keystone model using the RV test with Tax, RC, PC, and WP instruments. We report the results in Table 3.<sup>34</sup>

TABLE 3: Test Results

| Statistic | Instruments:  |                     |               |               |
|-----------|---------------|---------------------|---------------|---------------|
|           | (Preferred)   | (Other Instruments) |               |               |
|           | Tax<br>(1)    | RC<br>(2)           | PC<br>(3)     | WP<br>(4)     |
| $T^{RV}$  | 13.39         | 4.52                | 9.88          | 16.79         |
|           | ***           | ***                 | ***           | ***           |
| $F$       | 40.6          | 2.2                 | 15.6          | 1,591.6       |
|           | † † †   ^ ^ ^ | † † †   ^ ^ ^       | † † †   ^ ^ ^ | † † †   ^ ^ ^ |

The table reports, for each set of instruments, the RV test statistics  $T^{RV}$  and the effective  $F$ -statistic ([Duarte et al., 2024](#)) for testing Bertrand versus Keystone. A positive RV test statistic suggests a better fit of Keystone. The symbol \*\*\* indicates rejection of the null of equal fit 0.01 confidence level. The symbols † † † and ^ ^ ^ indicated that  $F$  is above the appropriate critical values for worst-case size below 0.075, and best-case power above 0.95, respectively. Both  $T^{RV}$  and the  $F$ -statistics account for two-step estimation error and clustering at the market level.  $n = 153,936$ .

In each column – corresponding to a set of instruments – the table reports the RV test statistic  $T^{RV}$  and the effective  $F$ -statistic ([Duarte et al., 2024](#)) for testing Bertrand versus

<sup>34</sup>Results are obtained with `pyRVtest` ([Duarte, Magnolfi, Sølvesten, Sullivan, and Tarascina, 2022](#)).

Keystone. A positive value of the RV test statistic above 1.96 indicates rejection of the null of equal fit in favor of better fit for Keystone at the 0.05 confidence level. All sets of instruments decisively reject the null in favor of Keystone. Given that we observe rejection, we are not concerned by low power, but we may still be concerned that these results are due to weak instruments leading to size distortions. For all instruments,  $F$  is above the critical value for the worst-case size of 0.075, suggesting size distortions are likely to be minimal.

Table 3 provides strong evidence that Keystone better describes conduct in this market. Across sets of instruments, each leveraging different features of the two models, the data consistently support Keystone as the superior model of conduct during our sample period, compared to Bertrand. In the next section, we explore how performing counterfactuals under the Keystone model has important implications for public policy.

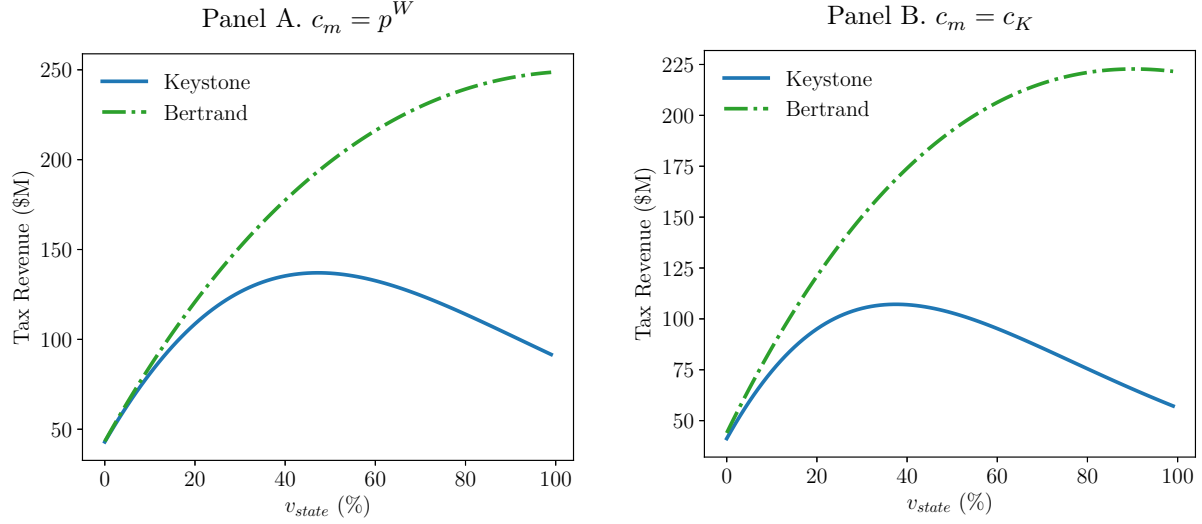
## 7.4 Counterfactuals

We perform three counterfactual exercises to illustrate the importance of learning conduct for optimal tax policy, focusing on the problem of setting the state-level ad valorem tax for the year 2016. While the government may have other considerations when setting taxes (e.g., curbing consumption to address externalities), we focus solely on revenue raised from the tax. Throughout, we hold market structure and local taxes fixed at their levels in the data.

The first two exercises highlight the importance of learning firm conduct for optimal tax policy counterfactuals. In both, we hold the primitives of demand and cost fixed and only vary the form of conduct. As policymakers might use wholesale price as a proxy for marginal cost, the first exercise sets marginal cost equal to the observed wholesale price. The second exercise imposes the marginal cost implied under Keystone, the model that best fits the data. In both exercises, we then simulate counterfactual prices, quantities, and government revenue at different levels of the state-wide ad valorem tax for each candidate model.

Our results are plotted in Figure 4. We find that the model of conduct has substantial implications for the Laffer curve when holding demand and cost fixed, in line with the conclusions from the recent literature (Miravete et al., 2018; Hollenbeck and Uetake, 2021; O’Connell and Smith, 2024). For both formulations of cost, the results are similar: the revenue-maximizing state-level tax rate is considerably lower under Keystone than Bertrand. In fact, when using the Keystone-implied costs in Panel B, the revenue maximizing state-level tax rate under Keystone is 38%, almost identical to the 37% rate the government adopted in 2015. The Bertrand model produces very different results; the associated Laffer curve lies well above the Keystone curve and is maximized at  $v_{state} = 90\%$  (Panel B).

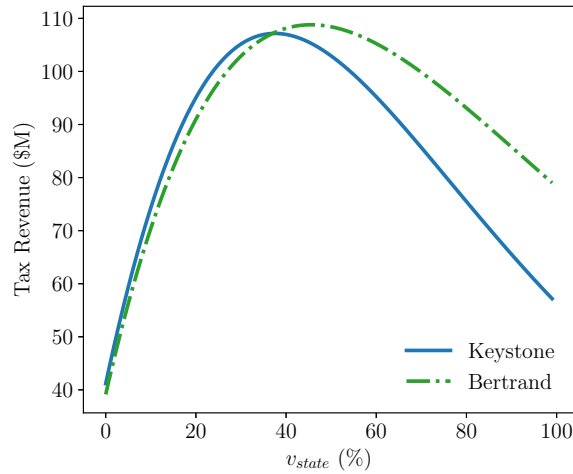
FIGURE 4: Laffer Curve in 2016 Fixing Cost Across Models



Panels A and B report, for 2016, the Laffer curve implied by different models of conduct for the state-level ad valorem tax (holding local tax rates fixed). Panel A imposes marginal cost equal to the wholesale price. Panel B imposes marginal cost equal to the estimates under Keystone pricing, the preferred model according to our testing results. We then simulate counterfactual prices, quantities, and government revenue at different levels of the tax for each candidate model.

In the third exercise, we consider the thought experiment of a policymaker who commits to a specific model of conduct to both recover costs and simulate the Laffer curve. To compare across assumptions on conduct, we obtain marginal cost estimates under each candidate model. We then use the model-specific cost estimates to simulate counterfactual prices, quantities, and government revenue at different levels of the state-wide ad valorem tax. Our results are plotted in Figure 5.

FIGURE 5: Laffer Curve in 2016 Under Model Implied Costs



We report for 2016, the Laffer curve implied by different models of conduct for the state-level ad valorem tax (holding local tax rates fixed). Using marginal cost estimates obtained under each model, we simulate counterfactual prices, quantities, and government revenue at different levels of the tax.

Since market outcomes are observed in the data, the Laffer curves for Bertrand and Keystone coincide at the tax rate in the data. Compared to Bertrand, the Keystone Laffer curve is maximized at a lower state-wide ad valorem tax rate (38% vs 46%). Compared to Figure 4, allowing costs to change under each model improves the model’s fit and its associated Laffer curve. However, there are quantitatively important consequences to adopting the wrong model. Setting tax at the level implied by Bertrand would lead to a loss of government revenue equal to \$2.06 million.

## 8 Conclusion

We discuss falsification of models of conduct in a general environment where researchers observe market outcomes for firms selling differentiated products. Our results highlight the economic features of different models that permit falsification, including the important role of cost pass-through. Thus, the falsifiable restriction in [Berry and Haile \(2014\)](#) generalizes the pass-through regression used in [Sumner \(1981\)](#) to learn firm conduct.

With our framework, we compare the relevance of various conduct instruments, such as rival cost shifters, product characteristics, and tax rates. In general, we find that different instruments target different features of the pass-through matrix. Furthermore, we show how our framework can guide instrument selection when counterfactual policy prediction is the primary objective. For counterfactuals designed to learn the revenue-maximizing tax, historical variation in the tax rate of interest provides the best instrument for learning conduct.

We demonstrate the usefulness of our framework in a set of simulations and an application to the Washington marijuana market. Appropriately-chosen ad valorem tax instruments conclude that a model of Keystone pricing best fits the data and that the state has chosen virtually the optimal tax rate. We anticipate that our results will offer a useful toolkit for researchers tackling new classes of models or instruments in a wide range of empirical settings.

## References

- ANDERSON, S., A. DE PALMA, AND B. KREIDER (2001): “The Efficiency of Indirect Taxes Under Imperfect Competition,” *Journal of Public Economics*, 81, 231–251.
- AUERBACH, A. (1985): “The Theory of Excess Burden and Optimal Taxation,” in *Handbook of Public Economics*, ed. by A. Auerbach and M. Feldstein, Amsterdam: Elsevier, vol. 1, chap. 2, 61–127.

- (2002): “Taxation and Corporate Financial Policy,” in *Handbook of Public Economics*, ed. by A. Auerbach and M. Feldstein, Amsterdam: Elsevier, vol. 3, chap. 19, 1251–1292.
- BACKUS, M., C. CONLON, AND M. SINKINSON (2021): “Common Ownership and Competition in the Ready-To-Eat Cereal Industry,” NBER Working Paper #28350.
- BERRY, S. AND P. HAILE (2014): “Identification in Differentiated Products Markets Using Market Level Data,” *Econometrica*, 82, 1749–1797.
- BRESNAHAN, T. (1982): “The Oligopoly Solution Concept Is Identified,” *Economics Letters*, 10, 87–92.
- BRUGUÉS, F. AND R. DE SIMONE (2024): “Taxation When Markets Are Not Competitive: Evidence From a Loan Tax,” Working Paper.
- BULOW, J. AND P. PFLEIDERER (1983): “A Note on the Effect of Cost Changes on Prices,” *Journal of Political Economy*, 91, 182–185.
- BUTTERS, R., D. SACKS, AND B. SEO (2022): “How Do National Firms Respond to Local Cost Shocks?” *American Economic Review*, 112, 1737–72.
- CONLON, C. AND J. GORTMAKER (2020): “Best Practices for Differentiated Products Demand Estimation with PyBLP,” *RAND Journal of Economics*, 51, 1108–1161.
- CORTS, K. (1999): “Conduct Parameters and the Measurement of Market Power,” *Journal of Econometrics*, 88, 227–250.
- DELLAVIGNA, S. AND M. GENTZKOW (2019): “Uniform Pricing in U.S. Retail Chains,” *Quarterly Journal of Economics*, 134, 2011–2084.
- DUARTE, M., L. MAGNOLFI, M. SØLVSTEN, AND C. SULLIVAN (2024): “Testing Firm Conduct,” *Quantitative Economics*, 15, 571–606.
- DUARTE, M., L. MAGNOLFI, M. SØLVSTEN, C. SULLIVAN, AND A. TARASCINA (2022): “pyRVtest: A Python Package for Testing Firm Conduct,” <https://github.com/anyatarascina/pyRVtest>.
- ESCUADERO, M. (2018): “Rule-of-Thumb Pricing: Retail Cannabis in Washington State,” Working Paper.
- GRIGOLON, L. AND F. VERBOVEN (2014): “Nested Logit or Random Coefficients Logit? A Comparison of Alternative Discrete Choice Models of Product Differentiation,” *Review of Economics and Statistics*, 96, 916–935.

- HANSEN, B., K. MILLER, AND C. WEBER (2022): “Vertical Integration and Production Inefficiency in the Presence of a Gross Receipts Tax,” *Journal of Public Economics*, 212, 104693.
- HOLLENBECK, B., S. HRISTAKEVA, AND K. UETAKE (2024): “Retailer Price Competition and Assortment Differentiation: Evidence from Entry Lotteries,” Working Paper.
- HOLLENBECK, B. AND K. UETAKE (2021): “Taxation and Market Power in the Legal Marijuana Industry,” *RAND Journal of Economics*, 52, 559–595.
- KROFT, K., J.-W. LALIBERTÉ, R. LEAL-VIZCAÍNO, AND M. NOTOWIDIGDO (2023): “Salience and Taxation with Imperfect Competition,” *Review of Economic Studies*, 91, 403–437.
- LAU, L. (1982): “On Identifying the Degree of Competitiveness from Industry Price and Output Data,” *Economics Letters*, 10, 93–99.
- LINDH, T. (1992): “The Inconsistency of Consistent Conjectures: Coming Back to Cournot,” *Journal of Economic Behavior & Organization*, 18, 69–90.
- MACKEY, A., N. MILLER, M. REMER, AND G. SHEU (2014): “Bias in Reduced-form Estimates of Pass-through,” *Economics Letters*, 123, 200–202.
- MAGNOLFI, L., D. QUINT, C. SULLIVAN, AND S. WALDFOGEL (2022a): “Falsifying Models of Firm Conduct,” Working Paper.
- (2022b): “Differentiated-Products Cournot Attributes Higher Markups than Bertrand-Nash,” *Economics Letters*, 219, 110804.
- MAKOWSKI, L. (1987): “Are ‘Rational Conjectures’ Rational?” *The Journal of Industrial Economics*, 36, 35–47.
- MIRAVETE, E., K. SEIM, AND J. THURK (2018): “Market Power and the Laffer Curve,” *Econometrica*, 86, 1651–1687.
- (2020): “One Markup to Rule Them All: Taxation by Liquor Pricing Regulation,” *American Economic Journal: Microeconomics*, 12, 1–41.
- (2024): “Elasticity and Curvature of Discrete Choice Demand Models,” Working Paper.
- O’CONNELL, M. AND K. SMITH (2024): “Optimal Sin Taxation and Market Power,” *American Economic Journal: Applied Economics*, forthcoming.
- PANZAR, J. AND J. ROSSE (1987): “Testing for ‘Monopoly’ Equilibrium,” *Journal of Industrial Economics*, 443–456.

- PLESS, J. AND A. A. VAN BENTHEM (2019): “Pass-Through as a Test for Market Power: An Application to Solar Subsidies,” *American Economic Journal: Applied Economics*, 11, 367–401.
- RAMSEY, F. (1927): “A Contribution to the Theory of Taxation,” *The Economic Journal*, 37, 47–61.
- RIVERS, D. AND Q. VUONG (2002): “Model Selection Tests for Nonlinear Dynamic Models,” *Econometrics Journal*, 5, 1–39.
- SLEMROD, J. (1990): “Optimal Taxation and Optimal Tax Systems,” *Journal of Economic Perspectives*, 4, 157–178.
- SUITS, D. AND R. MUSGRAVE (1953): “Ad Valorem and Unit Taxes Compared,” *The Quarterly Journal of Economics*, 67, 598–604.
- SULLIVAN, D. (1985): “Testing Hypotheses about Firm Behavior in the Cigarette Industry,” *Journal of Political Economy*, 93, 586–598.
- SUMNER, D. (1981): “Measurement of Monopoly Behavior: an Application to the Cigarette Industry,” *Journal of Political Economy*, 89, 1010–1019.
- WEYL, G. AND M. FABINGER (2013): “Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition,” *Journal of Political Economy*, 121, 528–583.