Nonparametric Identification of Demand without Exogenous Product Characteristics

Kirill Borusyak, Jiafeng Chen, Peter Hull, and Lihua Lei

ABSTRACT. We study the identification of differentiated product demand with exogenous supply-side instruments, allowing product characteristics to be endogenous. Past analyses have argued that exogenous characteristic-based instruments are essentially necessary given a sufficiently flexible demand model with a suitable index restriction. We show, however, that price counterfactuals are nonparametrically identified by recentered instruments—which combine exogenous shocks to prices with endogenous product characteristics—under a weaker index restriction and a new condition we term faithfulness. We argue that faithfulness, like the usual completeness condition for nonparametric identification with instruments, can be viewed as a technical requirement on the richness of identifying variation rather than a substantive economic restriction, and we show that it holds under a variety of non-nested conditions on either price-setting or the index.

JEL CODES. C14, C36, L13
KEYWORDS. Demand, nonparametric identification, recentered IV, faithfulness

*December, 2025. Borusyak: University of California, Berkeley, k.borusyak@berkeley.edu; Chen: Stanford University, jiafeng@stanford.edu; Hull: Brown University, peter_hull@brown.edu; Lei: Stanford University, lihualei@stanford.edu. We thank Nikhil Agarwal, Steve Berry, Jimbo Brand, and Phil Haile for helpful conversations. Google Gemini and OpenAI's ChatGPT contributed valuable insights.

1. Introduction

Classic parametric models of demand are often estimated with supply-side instruments for prices (e.g., Wright, 1928; Graddy, 1995; Blundell et al., 2012). Intuitively, exogenous variation in supply conditions across markets can trace out demand curves which can then be used to answer important counterfactuals—i.e., how quantities and welfare would change under hypothetical price changes. Modern analyses with more flexible demand models, however, often use other kinds of instruments: functions of the observed characteristics of competing products (e.g., Berry et al., 1995, 1999; Gandhi and Houde, 2019). It is well-known that such characteristic-based instruments can be invalid if product entry is strategic or if characteristic choice is otherwise endogenous (Berry et al., 1995; Petrin et al., 2022), and they can moreover reduce robustness to model misspecification (Andrews et al., 2025).

Can other instruments be found to avoid these issues? Are exogenous shocks to prices enough? In influential work, Berry and Haile (2014) argue that the answer is generally negative: i.e., that exogenous characteristic-based instruments are essentially necessary to flexibly estimate demand with market-level data. Specifically, they consider a nonparametric demand model of the form:

$$\sigma^{-1}(S, P) = X + \xi \equiv \delta, \tag{1}$$

where S and P are J-vectors of the observed quantity shares and prices of J goods in a market, X is a J-vector of an observed product characteristic, and ξ is an unobserved J-vector of demand shocks. The demand shocks and characteristics combine linearly in an index δ. The unknown $\sigma^{-1}(\cdot)$ is the inverse of a demand function $\sigma(d, p)$ giving the market shares that would be observed if δ were set to d and P were set to p.[1]

Berry and Haile (2014) argue that observing a J-vector of exogenous price instruments Z, which shift S through P while being unrelated to ξ, is not enough to identify $\sigma(\cdot)$ or even to identify price counterfactuals (e.g., functions of $\partial\sigma/\partial p$).[2] Intuitively, the inverse demand function $\sigma^{-1}(\cdot)$ depends on S and P separately but Z shifts both simultaneously; to disentangle price effects we would thus seem to need other instruments that shift market shares given prices.[3] The characteristic X that enters the

[1] $\sigma(\cdot)$ may also depend on other observable characteristics $\tilde{X}$, which we suppress here for exposition.

[2] This latter claim is most directly stated in a review, Berry and Haile (2021, p. 6): "having valid instruments for all J prices will not generally suffice for identification of [...] the *ceteris paribus* effects of price changes."

[3] Berry and Haile (2014) also consider a model in which P enters the index and is excluded from $\sigma^{-1}(\cdot)$, concluding that price instruments suffice for identification, though again assuming that

index uniquely fits this bill, so long as it is also exogenous. All functions of (X, Z) can then serve as instruments and identify $\sigma(\cdot)$ provided they induce enough variation in (S, P)—a technical condition known as completeness (Newey and Powell, 2003).

We reexamine this setting and arrive at a different, more optimistic conclusion: exogenous shocks to prices can suffice for nonparametric identification of price counterfactuals. Intuitively, even when characteristics are endogenous, functions of X and an as-good-as-randomly assigned Z that are mean-zero given X—what Borusyak and Hull (2023) call recentered instruments—are still valid and have identifying power. In fact, they exhaust the exogenous variation in $Z \mid X$ and can rule out some candidate demand functions $\check{\sigma}(\cdot)$ through a simple test: if $\check{\delta} = \check{\sigma}^{-1}(S, P)$ correlates with a recentered instrument, then $\check{\sigma}(\cdot)$ cannot be the true demand function.

We introduce a new technical condition—*faithfulness*—under which all price counterfactuals are identified by the candidate demand functions that survive this test. Specifically, we show that, under faithfulness, any surviving $\check{\sigma}^{-1}(\cdot)$ equals the true $\sigma^{-1}(\cdot)$ up to some unknown transformation which is inconsequential for price counterfactuals. Naturally, this argument does not fully identify the demand function, as we cannot learn counterfactuals that change the endogenous characteristics. But for identifying price counterfactuals, we show the other characteristics need not be exogenous and that conventional characteristic-based instruments may be unnecessary.[4]

The new faithfulness condition intuitively requires the price instruments to induce enough exogenous variation to make all causal effects of P detectable. Specifically, it says that functions of the form $H(\delta, P)$ cannot be mean-independent of Z given X unless they are constant in P; otherwise, there are causal effects of P on $H(\delta, P)$ that are undetected by Z. This condition is useful because each candidate $\check{\sigma}^{-1}(\cdot)$ can be written as such a function: $\check{\sigma}^{-1}(S, P) = \check{\sigma}^{-1}(\sigma(\delta, P), P) \equiv \check{H}(\delta, P)$. Thus, if faithfulness holds, any $\check{\sigma}^{-1}(S, P)$ that survive the recentered instrument test must be a transformation of the true $\sigma^{-1}(S, P) = \delta$: i.e., $\check{H}(\delta, P) = \check{H}(\delta) = \check{H}(\sigma^{-1}(S, P))$.

Like completeness, faithfulness is best seen as a technical restriction on the richness of identifying variation rather than a substantive economic assumption. The two conditions are indeed similar: completeness says that functions of the form $H(S, P)$ cannot be mean-independent of (X, Z) unless they are constant in (S, P). Faithfulness

characteristics are exogenous. Borusyak et al. (2025) show that recentered instruments identify price counterfactuals in this case, allowing endogenous characteristics.

[4] As we discuss below, allowing characteristics to be endogenous further relaxes the index restriction in the main text of Berry and Haile (2014) by allowing δ to be an arbitrary function of X and ξ.

focuses on the richness of exogenous price variation given potentially endogenous characteristics, in line with the goal of identifying price counterfactuals.

We establish several non-nested conditions under which the two technical conditions are quite close. One simple but instructive case is when δ and X have finite support with the same number of values; we show that completeness implies faithfulness in this case, suggesting that the gap between the two assumptions can only arise in the presence of infinite-dimensional objects. We also show that when prices are exogenous (i.e., $P = Z$)—a setting that past analyses have characterized as surprisingly difficult for identification (Berry and Haile, 2021)—faithfulness follows from completeness.

We further develop several conditions on either pricing or the utility index under which faithfulness also follows from completeness. On the pricing side, it is sufficient that either *(i)* X and Z combine in a nonparametric index λ such that P is independent of (X, Z) given (λ, δ) or *(ii)* the derivatives of P with respect to Z satisfy a particular separability condition. Both conditions are compatible with firms engaging in Bertrand–Nash pricing with certain forms of marginal costs. On the δ index side, we show that faithfulness follows from completeness when $\delta \mid X$ can be transformed to follow a location-scale model satisfying certain smoothness conditions—regardless of how prices are determined. Importantly, researchers do not need to take a stand on the specific sufficient condition, as the identification argument is the same so long as faithfulness holds. Together, our sample of sufficient conditions shows that faithfulness can hold without strong economic or statistical restrictions.

Overall, our results formalize a straightforward intuition for the amount of exogenous variation needed to identify counterfactuals in demand models. The model specifies potential outcomes in two "treatments," prices and characteristics. If the causal effects of both treatments are of interest—i.e., if we want to identify the entire $\sigma(\cdot)$ function—then it is not surprising that exogenous variation in both X and P is needed, leading to the familiar "$2J$" instrument requirement in Berry and Haile (2014). However, if we are only interested in the causal effects in prices, then it is intuitive that only exogenous variation in P is needed. Indeed, our results show that identifying price counterfactuals does not require separating out how s and p enter $\sigma^{-1}(s, p)$; it only requires finding candidate $\check{\sigma}^{-1}(\cdot)$ for which $p \mapsto \check{\sigma}^{-1}(\sigma(\delta, p), p)$ is constant, which can be checked with exogenous variation in P only. Faithfulness makes this exclusion of P testable via recentered instrument moment conditions.

Our results also resolve seemingly conflicting recommendations for empirical researchers. In parametric settings, Borusyak et al. (2025) and Andrews et al. (2025)

recommend practitioners use recentered instruments to make their demand estimates more robust to endogeneity and other misspecification concerns. However, given prevailing intuition that exogenous characteristics are essentially necessary for nonparametric identification, one might worry that such robustness is tied to some implicit parametric assumptions. By showing J-dimensional exogenous variation can be nonparametrically sufficient, our results reassure practitioners that this is not the case.

This paper contributes to two main literatures. First and most centrally, we add to a prominent literature on the identification of differentiated product demand with market-level data starting with Berry (1994) and Berry et al. (1995, 1999). Most closely related is the nonparametric analysis of Berry and Haile (2014). We depart from much of this literature by allowing product characteristics to be endogenous and by focusing on the identification of price counterfactuals, what Berry and Haile (2024) term conditional demand.[5] Other papers in this spirit but focusing on the estimation of parametric models include Ackerberg and Crawford (2009) and Borusyak et al. (2025), with the latter proposing recentered instruments for estimation.

The model we study imposes an index restriction in the spirit of Berry and Haile (2014), albeit a substantially weaker one. Thus, it is still restrictive relative to the fully-general potential outcomes model that Angrist et al. (2000) consider for estimating the demand for a single good by linear instrumental variables (IV). The index restriction is important for identifying market-specific price counterfactuals, which are important for many economic questions and generally not given by the local average demand elasticities that linear IV identifies (Berry and Haile, 2021; Chen, 2025).

Second, our results add to a classic literature on nonparametric identification with instrumental variables, including Brown and Matzkin (1998), Newey et al. (1999), Newey and Powell (2003), Altonji and Matzkin (2005), Chernozhukov et al. (2007), Benkard and Berry (2006), Chiappori et al. (2015), Imbens and Newey (2009), Torgovitsky (2015), and D'Haultfœuille and Février (2015). Beyond demand, our results can be restated to apply to triangular models in which unobserved heterogeneity enters the second stage through an index with a potentially endogenous covariate.[6] Our faithfulness condition appears new and may prove useful for identifying other nonparametric models with this structure.[7]

[5] Berry and Haile (2024) study nonparametric identification when the kind of market-level data we consider here is augmented with micro data linking individual consumers' characteristics and choices.

[6] Specifically, our results apply to the model characterized by $Y = g(W_1, \delta(W_2, \xi))$ and $W_1 = h(Z, W_2, \omega)$ with $Z \perp\!\!\!\perp (\xi, \omega) \mid W_2$ and $(Y, W_1, \delta, W_2) \in \mathbb{R}^J$.

[7] Our faithfulness condition derives its name from and relates conceptually to a condition in the causal discovery literature (Spirtes et al., 2000), which considers whether a joint distribution of variables is

The rest of this paper is organized as follows. Section 2 develops the main identification results. Section 3 presents sufficient conditions for the new faithfulness condition. Section 4 concludes. For clarity, we suppress various technical details in the main text (e.g., regularity conditions on measurability, existence of moments, and support) relegating detailed statements and proofs to Appendix A.

2. Theory

2.1. **Setup.** Consider a demand system in a given market with J products and quantity shares (or other quantity measures) $S \in \mathbb{R}^J$ determined by

$$S = \sigma(\delta(\overline{X}, \xi), \tilde{X}, P), \tag{2}$$

with prices $P \in \mathbb{R}^J$, observed characteristics $\overline{X} = (X, \tilde{X}) \in \mathbb{R}^J \times \mathcal{X}$, and unobserved shocks $\xi \in \Xi$ incorporating latent consumer tastes or unobserved product characteristics.[8] Following Berry and Haile (2014), Equation (2) imposes an index restriction in which the demand shocks and characteristics enter together as $\delta(\overline{X}, \xi) \in \mathbb{R}^J$, with the "special characteristic" X otherwise excluded and influencing S only through δ.[9]

We are interested in identifying price counterfactuals: the quantities $\sigma(\delta(\overline{X}, \xi), \tilde{X}, p')$ that would have been observed had prices been set to some p'. To do so, we assume a set of price instruments $Z \in \mathcal{Z}$ is observed in addition to $(S, \overline{X}, P)$. The identification challenge stems from the unobservability of ξ along with the unknown σ and δ.

We follow Berry and Haile (2014) by assuming the demand function is invertible:

Assumption 1. *For every $(\tilde{x}, p)$, the map $d \mapsto \sigma(d, \tilde{x}, p)$ is invertible.*

Invertibility follows from a weak "connected substitutes" condition (Berry et al., 2013).

Berry and Haile (2014) consider Equation (2) with the additional assumptions of *(i)* exclusion of $\tilde{X}$ from the utility index, $\delta(\overline{X}, \xi) = \delta(X, \xi)$, *(ii)* a linear index, $\delta(X, \xi) = X + \xi$ with $\xi \in \mathbb{R}^J$, and *(iii)* mean-independence of the unobserved demand shocks from the characteristics and instruments: $\mathbb{E}[\xi \mid \overline{X}, Z] = 0$. The last assumption captures the exogeneity of $\overline{X}$ in addition to Z. We combine these three assumptions

"faithful" to an underlying causal graph; Appendix B.1 details this connection. Appendix B.2 shows that some nonparametric identification results in Imbens and Newey (2009), Torgovitsky (2015), and D'Haultfœuille and Février (2015) can be interpreted as verifying an analogue of faithfulness.

[8]Focusing on identification, we suppress subscripts for specific markets in a sample.

[9]Note that the meaning of the index δ differs from that in the literature on parametric demand models following Berry et al. (1995) where it usually denotes the mean utility vector over products. Most importantly, here δ does not include the effect of price on demand.

in a slightly weaker form, without imposing *(i)* or *(ii)*:

$$\mathbb{E}[\delta(\overline{X}, \xi) - X \mid \overline{X}, Z] = 0. \tag{3}$$

This condition yields identification of σ under a standard completeness condition (Newey and Powell, 2003; Ai and Chen, 2003; Andrews, 2011; D'Haultfoeuille, 2011; Miao et al., 2018):

Assumption 2. *The distribution of* $(S, P, \tilde{X}) \mid (X, Z, \tilde{X})$ *is complete: For all* h *with* $\mathbb{E}[\|h(S, P, \tilde{X})\|] < \infty$, $\mathbb{E}[h(S, P, \tilde{X}) \mid X, Z, \tilde{X}] = 0 \implies h(S, P, \tilde{X}) = 0$.

Identification of σ follows from the fact that, under Assumption 1 and Equation (3),

$$X = \mathbb{E}[\sigma^{-1}(S, P, \tilde{X}) \mid \overline{X}, Z], \tag{4}$$

while Assumption 2 ensures the solution to this integral equation in σ^{-1} is unique.

Our analysis works under a partial relaxation of Equation (3):

Assumption 3. $Z \perp\!\!\!\perp \xi \mid \overline{X}$. *Equivalently,* $Z \perp\!\!\!\perp \delta \mid \overline{X}$.

This restriction strengthens the mean independence in (3) to the full independence of Z given $\overline{X}$ but crucially drops the exogeneity of $\overline{X}$. Hence, under Assumption 3,

$$\mathbb{E}[\delta(\overline{X}, \xi) \mid \overline{X}, Z] = \mathbb{E}[\delta(\overline{X}, \xi) \mid \overline{X}] \equiv k_0(\overline{X}), \tag{5}$$

for some unknown $k_0(\overline{X})$ which need not equal X. Assumption 3 can be motivated by viewing Z as a set of supply-side shocks drawn in some true or natural experiment, after product characteristics are determined (Borusyak et al., 2025).[10]

Our relaxation of Equation (3) is motivated by three related misspecification concerns. First, suppose the functional form $\delta(\overline{X}, \xi) = X + \xi$ is correctly specified but that characteristics X are chosen strategically by firms with partial knowledge of demand conditions as captured by ξ. Then $\mathbb{E}[\xi \mid \overline{X}] = 0$ is unlikely. Second, ξ may include physical product characteristics chosen by firms along with $\overline{X}$ but unobserved by the econometrician. Again, there is little reason they would be uncorrelated with $\overline{X}$. Finally, suppose the characteristics $\overline{X}$ are indeed exogenous (e.g., fully independent of ξ) but that the functional form of δ is not linear or does not exclude $\tilde{X}$: $\delta(\overline{X}, \xi) \neq X + \xi$. Then (3) need not hold, and model-based price counterfactuals relying on (3) may be incorrect (Andrews et al., 2025). Relaxing the linear index

[10] For example, in the U.S. automobile market, Z might contain tariff or exchange rate shocks in cars' countries of production that are excluded from demand and drawn randomly after cars are designed but before prices are set. In this case $Z \perp\!\!\!\perp (\xi, \overline{X})$, which implies Assumption 3.

functional form substantively weakens restrictions on the heterogeneity of causal effects of $\overline{X}$ on S (Chen, 2025). This relaxation also allows ξ to be multi-dimensional for each product, in contrast to most models in the literature.[11]

Equation (5) motivates our identification strategy. Under Assumption 3, the true inverse demand function is mean-independent of Z conditional on $\overline{X}$:

$$\begin{aligned}\mathbb{E}[\sigma^{-1}(S, P, \tilde{X}) \mid \overline{X}, Z] &= \mathbb{E}[\delta(\overline{X}, \xi) \mid \overline{X}, Z] \\ &= \mathbb{E}[\delta(\overline{X}, \xi) \mid \overline{X}] = \mathbb{E}[\sigma^{-1}(S, P, \tilde{X}) \mid \overline{X}].\end{aligned} \tag{6}$$

We can thus rule out candidate functions $\check{\sigma}^{-1}$ that violate this conditional moment restriction. A natural way is by using recentered instruments, of the form

$$R(\overline{X}, Z) - \mathbb{E}\left[R(\overline{X}, Z) \mid \overline{X}\right] \quad \text{for some } R.$$

We show in Lemma B.1 that the set of $\check{\sigma}^{-1}$ which are orthogonal to all such instruments is exactly the set satisfying the conditional moment restriction (6), building on earlier results in Borusyak and Hull (2023) and Borusyak et al. (2025).

We show below that any such $\check{\sigma}^{-1}$ yields correct price counterfactuals under a faithfulness condition. Before proceeding, however, we simplify some notation. Our identification results fully condition on the characteristics $\tilde{X}$ so, for simplicity, we drop them from the arguments of $\sigma^{-1}(S, P)$ and replace $\overline{X}$ with X. We also define

$$\Theta_I \equiv \left\{h(s,p),\ \mathbb{R}^J\text{-valued and invertible in } s\colon \mathbb{E}[h(S,P) \mid X, Z] = k(X) \text{ for some } k\right\}.$$

as a (potentially non-sharp) identified set: the set of candidate $\check{\sigma}^{-1}$ that are orthogonal to all recentered instruments. Θ_I is nonempty since all transformations of $\sigma^{-1}(S, P)$ are independent of Z given X under Assumption 3. Lastly, we shorthand $\delta(X, \xi) = \delta$.

2.2. **Main results.** Our first result shows that, for identifying price counterfactuals, it is sufficient for σ^{-1} to be identified up to some transformation T:

Lemma 1. *Suppose all $\check{\sigma}^{-1} \in \Theta_I$ have the form $\check{\sigma}^{-1}(S, P) = T(\sigma^{-1}(S, P))$ for some $T : \mathbb{R}^J \to \mathbb{R}^J$. Then price counterfactuals are identified.*

Intuitively, counterfactual quantities given potential prices p' under candidate model $\check{\sigma}^{-1} = T(\sigma^{-1})$ are computed from the observed (S, P) as:

$$\check{\sigma}(\check{\sigma}^{-1}(S, P), p') = \sigma(T^{-1}(T(\sigma^{-1}(S, P))), p') = \sigma(\delta, p').$$

[11]For instance, Appendix B of Berry and Haile (2014) contains a nonseparable model in which $\delta = \delta(X, \xi)$, but restricts the map coordinate-wise: $\delta_j = \delta_j(X_j, \xi_j)$ with the function increasing in ξ_j. This requires ξ to be of dimension J, in addition to other restrictions that we remove.

The transformation T cancels, so it does not affect the counterfactual $\sigma(\delta, p')$.[12]

We next show how the condition in Lemma 1 can be satisfied. Consider:

Assumption 4 (Faithfulness). *The distribution of $(\delta, P) \mid (X, Z)$ is* faithful*: for all $\mathbb{R}^J$-valued $H(\delta, P)$, if $\mathbb{E}[H(\delta, P) \mid X, Z] = k(X)$ does not depend on Z then H does not depend on P, i.e., $H(\delta, P) = H(\delta)$.*

To see that this faithfulness condition is sufficient, write

$$\check{\sigma}^{-1}(S, P) = \check{\sigma}^{-1}(\sigma(\delta, P), P) \equiv \check{H}(\delta, P)$$

and note that when $\check{\sigma}^{-1} \in \Theta_I$:

$$\mathbb{E}[\check{H}(\delta, P) \mid X, Z] = \mathbb{E}[\check{\sigma}^{-1}(S, P) \mid X, Z] = k(X) \text{ for some } k.$$

Under faithfulness, this means that $\check{\sigma}^{-1}(S, P) = \check{H}(\delta, P) = \check{H}(\delta) = \check{H}(\sigma^{-1}(S, P))$ is a transformation of the true inverse demand function, such that all price counterfactuals are identified by Lemma 1. We summarize this argument in the following result:

Proposition 1. *Under Assumptions 1, 3, and 4, price counterfactuals are identified.*

Faithfulness enables identification from (6) much like completeness enables identification from (4). Indeed, note that completeness is not needed for Proposition 1 once faithfulness is imposed. In the nonparametric IV literature, completeness is often treated as a technical condition rather than a substantive restriction; we argue that faithfulness should be treated similarly. Both conditions are about whether certain changes are detectable through conditional expectations. Completeness of $(\delta, P) \mid (X, Z)$ means that $\mathbb{E}[H(\delta, P) \mid X, Z] = 0$ implies $H = 0$ (Assumption 2).[13] In other words, all changes in (δ, P)-space are detectable even when we project to (X, Z)-space. In this sense, the conditional expectation "faithfully reflects" changes in (δ, P) space. In a similar vein, faithfulness requires that the conditional expectation faithfully represents changes in P—in the sense that any change in P is detectable via changes in Z after projecting onto the (X, Z)-space.

Because of this similarity, all the sufficient conditions that follow verify faithfulness from completeness plus additional restrictions. Some additional conditions are needed, since faithfulness does not always follow from completeness; we construct a counterexample in Appendix B.3. That said, we have found such examples hard to construct—suggesting faithfulness is not much more restrictive than completeness.

[12] Note that T is implicitly invertible because Θ_I only includes invertible $\check{\sigma}$ candidates.

[13] Completeness of $(S, P) \mid (X, Z)$ is equivalent to completeness of $(\delta, P) \mid (X, Z)$ since σ is invertible.

Here we give two simple but instructive sufficient conditions for faithfulness, with other more involved conditions given below. First, we show that if δ and X both take the same finite number of values then faithfulness is a corollary of completeness. The reason is that any function $k(X)$ can be represented as a projection of some function $H(\delta)$ to (X, Z)-space in this case. Completeness then forces $H(\delta, P) = H(\delta)$.

Proposition 2. *Fix $M \in \mathbb{N}$. Assume the support of δ and X are both finite sets of size M. Under Assumptions 2 and 3, Assumption 4 holds.*

This result suggests that faithfulness is only a stronger condition than completeness in the presence of infinite-dimensional objects.

Second, we show that completeness implies faithfulness when prices are as-good-as-randomly assigned. While exogenous prices is unlikely in most applications, it serves as an illuminating baseline case which Berry and Haile (2021) call "surprisingly difficult." This case also informs the intuition for why exogenous variation in P is sufficient, which we discuss at length below.

Proposition 3. *Suppose $P = Z$.*[14] *Then Assumptions 2 and 3 imply Assumption 4.*

To show this, fix some value p_0. Given $H(\delta, P)$ with $\mathbb{E}[H \mid P, X] = k(X)$,

$$0 = \mathbb{E}[H(\delta, p) \mid P = p, X = x] - \mathbb{E}[H(\delta, p_0) \mid P = p_0, X = x] \tag{7}$$

$$= \mathbb{E}[H(\delta, p) - H(\delta, p_0) \mid X = x] \tag{8}$$

$$\implies 0 = H(\delta, p) - H(\delta, p_0) \text{ by Assumption 2}, \tag{9}$$

where the first line uses that Z does not enter $k(X)$ and the second line uses Assumption 3. Thus $H(\delta, p) = H(\delta, p_0)$ does not depend on p, satisfying faithfulness.

2.3. **Discussion: what exogenous variation is needed?** Our results contrast with prevailing intuition, since Berry and Haile (2014), that exogenous variation in X is needed to flexibly identify demand or price counterfactuals. The standard argument is based on observing the inverted structural equation

$$X = \sigma^{-1}(S, P) - \xi.$$

There are $2J$ endogenous variables (S, P) on the right-hand side of this equation, suggesting a need for $2J$ instruments: J instruments "for the shares" and J instruments "for the prices." Since X is excluded from the right-hand side, nonlinear transformations of it can serve as instrument when characteristics are exogenous. The role of an

[14]Note also in this case, under Assumption 3, completeness of $(\delta, P) \mid (X, P)$ is equivalent to completeness of $\delta \mid X$.

exogenous X, then, is not to generate variation in prices but to generate variation in shares conditional on prices so that one can disentangle the two arguments of σ^{-1}.

This argument in inverse-demand space concurs with a straightforward intuition: if we would like to learn $\sigma(\cdot)$, a causal relationship between S and two J-dimensional "treatments" (X, P), then clearly we would need exogenous variation in both treatments. But this leaves open the possibility that less exogenous variation is needed if we are only interested in the causal effects of prices—i.e., price counterfactuals.

Our first key insight is that, for price counterfactuals, it suffices to identify σ^{-1} up to a transformation. From the perspective of causal effects, this requirement is equivalent to finding *some* outcome $H \equiv h(S, P)$, invertible in S, such that prices have no effect on it. For any such h, counterfactuals at p are correctly predicted by $h^{-1}(h(S, P), p)$. Importantly, we are only interested in whether P has causal effects on H and not in disentangling the effects through S and P—sidestepping the need for separate variation in S given P.

To verify whether P has causal effects on H, one needs exogenous variation in prices. In the simplest case of exogenous prices, $P = Z$, we can directly trace the effect on H *on average*, via the cross-price comparisons in Equations (7) and (8). This, however, is not enough: while the index restriction implies that the effect of P on S and therefore on H is fully governed by δ, the δ is unobserved. Zero average price effects—averaging across markets with different δ's—does not by itself imply zero price effects for markets with a particular δ. This is reflected in the argument for Proposition 3: Equation (9) does not follow from Equation (8) without assumptions.

This identification gap is filled by leveraging variation in X, which needs to be rich enough—but need not be exogenous.[15] For each stratum $X = x$, exogenous price variation yields a conditional average causal effect of price on H, as in Equation (8). With strata rich enough to span the variation in δ, zero conditional average effects for each $X = x$ guarantees that there cannot be non-zero effects for any $\delta = d$, i.e. Equation (9). The richness of the variation in X is formalized by the completeness of the $\delta \mid X$ distribution, which for exogenous prices implies faithfulness (Proposition 3). The role of X in identification here is thus not in "instrumenting for shares," but in proxying for the unobserved δ.

The general version of faithfulness captures the same intuition with endogenous prices. In that case, the condition requires both the richness of the price variation

[15] Section 3 in Chen (2025) shows that many counterfactual predictions from structural models analogously extrapolate from a lack of average causal effects to a total lack of causal effects.

induced by Z and the richness of the variation in δ induced by X. We next provide three non-nested sufficient conditions for the general case, further evidencing that faithfulness does not require strong economic or statistical restrictions.

3. **Sufficient conditions**

3.1. **On price-setting.** We start from two sufficient conditions that restrict how price depends on the observables and unobservables of the model. Both conditions extend the case of exogenous prices. For some function f, write

$$P = f(X, Z, \delta, \tilde{\omega}), \qquad \tilde{\omega} \perp\!\!\!\perp (X, Z, \delta). \tag{10}$$

Here $\tilde{\omega}$ is an unobservable of arbitrary dimension that captures residual variation in P that is independent of (X, Z, δ). So far, Equation (10) is without loss of generality.[16] In what follows we place different restrictions on f. At the end of this subsection, we show that these restrictions can be restated as similar conditions on marginal costs under Bertrand–Nash pricing.

The first sufficient condition is that X and Z enter price only through the utility index δ and an index $\lambda(X, Z) = \lambda$ that is invertible in Z:

Assumption 5. *$P \perp\!\!\!\perp (X, Z) \mid (\lambda(X, Z), \delta)$, for some $\lambda(x, z)$ that is invertible in z. Equivalently, in (10), $f(X, Z, \delta, \tilde{\omega}) = f(\lambda(X, Z), \delta, \tilde{\omega})$.*

This assumption is satisfied, in particular, when X enters price in (10) only through δ—paralleling how it enters $\sigma(\cdot)$. Under this index restriction for price, Assumption 5 is satisfied with $\lambda(X, Z) = Z$. By further setting $Z = P$, this special case also nests exogenous prices and generates Proposition 3 as a corollary. In general we have:

Proposition 4. *Under Assumptions 2, 3, and 5, Assumption 4 holds.*

To see how this result follows, note that for any realization λ_0 of λ:[17]

$$\begin{aligned}
k(X) &= \mathbb{E}[H(\delta, P) \mid X, Z] = \mathbb{E}[\mathbb{E}[H(\delta, P) \mid \delta, X, Z, \lambda] \mid X, Z] \\
&= \mathbb{E}[\mathbb{E}[H(\delta, P) \mid \delta, \lambda] \mid X, Z] && \text{(Assumption 5)} \\
&= \mathbb{E}[\mathbb{E}[H(\delta, P) \mid \delta, \lambda = \lambda_0] \mid X, Z = \lambda^{-1}(X, \lambda_0)] && (Z \text{ does not enter } k(X)) \\
&= \mathbb{E}[\mathbb{E}[H(\delta, P) \mid \delta, \lambda = \lambda_0] \mid X] && \text{(Assumption 3)}
\end{aligned}$$

[16]Note that f is not a structural function because of the parametrization of $\tilde{\omega}$, but it is consistent with any structural formulation. Indeed, for any structural shock ω possibly correlated with (X, Z, δ) and $P = \tilde{f}(X, Z, \delta, \omega)$, one can represent $\omega = f_\omega(X, Z, \delta, \tilde{\omega})$ and $P = \tilde{f}(X, Z, \delta, f_\omega(X, Z, \delta, \tilde{\omega}))$.

[17]When λ is continuously distributed, conditioning on a realization of it requires measure-theoretic care. See Proposition A.4 for a formal argument.

$$= \mathbb{E}[\mathbb{E}[H(\delta, P) \mid \delta, \lambda = \lambda_0] \mid X, Z]. \qquad \text{(Assumption 3)}$$

Hence for $H(\delta) = \mathbb{E}[H(\delta, P) \mid \delta, \lambda = \lambda_0]$ we have $\mathbb{E}[H(\delta) \mid X, Z] = k(X)$. Finally, by completeness (Assumption 2), $H(\delta, P) = H(\delta)$.

The second sufficient condition instead imposes a separability condition on the derivatives of price with respect to Z:

Assumption 6. *In (10), f is continuously differentiable in Z with Jacobian $D_z f$, which satisfies the following separability condition: for measurable functions A, B,*

$$\underbrace{D_z f(X, Z, \delta, \tilde{\omega})}_{J \times d_z} = \underbrace{A(f(X, Z, \delta, \tilde{\omega}), \delta)}_{J \times J} \cdot \underbrace{B(X, Z)}_{J \times d_z} \tag{11}$$

where $A(f(X, \delta, Z, \tilde{\omega}), \delta)$ and $B(X, Z)$ are full row rank with $\dim(Z) \equiv d_z \geq J$.

Assumption 6 holds, in particular, when

$$P = f_0\bigg(f_1(X, Z) + f_2(X, \delta, \tilde{\omega}), \delta\bigg) \tag{12}$$

(with regularity conditions on f_0 and f_1 given in Lemma A.3). The restriction in Equation (12) is that Z enters P through an index $f_1(X, Z) + f_2(X, \delta\tilde{\omega})$, which is a form of separability between observed cost shifters Z and unobserved ones, $\tilde{\omega}$. Clearly, it is satisfied when $P = Z$.

Proposition 5. *Under Assumptions 1 to 3 and 6, Assumption 4 holds.*

The logic for this result is as follows: under Equation (11), given any differentiable $H(\delta, P)$ with $H_p(\delta, P) \equiv \frac{\partial H}{\partial P}$,[18] we have

$$\begin{aligned}
0_{J \times d_z} &= \frac{\partial}{\partial z}\mathbb{E}[H(\delta, P) \mid X, Z = z] = \frac{\partial}{\partial z}\mathbb{E}[H(\delta, f(X, Z, \delta, \omega)) \mid X, Z = z] \\
&= \mathbb{E}\left[H_p(\delta, P)A(P, \delta) \mid X, Z = z\right] B(X, z) && \text{(Assumption 6)} \\
\implies 0_{J \times J} &= \mathbb{E}\left[H_p(\delta, P)A(P, \delta) \mid X, Z\right] && \text{(B is full-rank)} \\
\implies 0_{J \times J} &= H_p(\delta, P)A(P, \delta) && \text{(Assumption 2)} \\
\implies 0_J &= H_p(\delta, P) && \text{(A is nonsingular)}
\end{aligned}$$

By the fundamental theorem of calculus, $H_p(\delta, P) = 0$ implies $H(\delta, P) = H(\delta)$.

We emphasize that Assumptions 5 and 6 are non-nested. The former is more flexible on how Z and $\tilde{\omega}$ can interact: e.g., $P = f(Z, \tilde{\omega})$ always satisfies Assumption 5 but not

[18]Because we assume H is differentiable here, we have to modify faithfulness to restrict only differentiable functions. These complications are resolved in Proposition A.5, which redefines faithfulness and ensures that differentiation is exchangeable with expectation.

necessarily Assumption 6. The latter is more flexible on how X interacts with $\tilde{\omega}$: e.g., $P = Z + f(X, \tilde{\omega})$ always satisfies Assumption 6 but not necessarily Assumption 5.

Assumptions 5 and 6 can be economically grounded with more primitive conditions on marginal costs. Under Bertrand–Nash pricing with constant marginal costs that are represented without loss of generality as $c(X, Z, \delta, \tilde{\omega}) \equiv C \in \mathbb{R}^J$, one can always write the equilibrium prices as

$$P = g(C, \delta) \tag{13}$$

for some function g (see Appendix B.5). Thus, if one assumes that X and Z enter marginal costs via the index λ, i.e., $C = c(\lambda(X, Z), \delta, \tilde{\omega})$, then Assumption 5 follows. Similarly, Assumption 6 holds if $C = f_0(f_1(Z, X) + f_2(\delta, X, \tilde{\omega}), \delta)$ with g and f_0, f_1 satisfying certain regularity conditions.

3.2. **On the δ index.** Our third sufficient condition exploits smoothness restrictions on the conditional distribution of $\delta \mid X$ that might be imposed *a priori*. While this condition is high-level, it is used only to demonstrate existence of such restrictions—not necessarily as a recommended modeling restriction. In particular, suppose $\delta \mid X$ can be transformed into a location-scale model of the form:

$$a(\delta) = b(X) + \Sigma(X)\epsilon, \quad \epsilon \sim q(\cdot), \quad \epsilon \perp\!\!\!\perp X, \tag{14}$$

for some invertible $a(\cdot)$ and continuously distributed ϵ with density $q(\cdot)$; here ϵ can be seen as reparametrizing the component of ξ that is independent of X.

We consider the following assumptions on $a(\cdot)$, $b(\cdot)$, $\Sigma(\cdot)$, and $q(\cdot)$:

Assumption 7. *Fix some integers $s > J + 1$ and $M = J + 2s + 1$. Let*

$$\mathbb{K}^s = \left\{u : \mathbb{R}^J \to \mathbb{R}^J : u(x) = Ax + r(x), A \in \mathbb{R}^{J\times J}, r \in \mathbb{W}^{s,\infty}(\mathbb{R}^J)\right\}$$

where $\mathbb{W}^{s,p}(\mathbb{R}^J)$ is a Bessel potential space of smoothness parameter s and integrability parameter p, defined in (A.8). Let $\mathcal{K} = \mathbb{K}^M$. Assume

(1) a, b in (14) are invertible.

(2) (Density smoothness for ϵ) The Fourier transform of q and its derivatives have bounded tails obeying Assumption A.7(2).

(3) (Limited heteroskedasticity) $\Sigma(x)$ is uniformly close to some fixed Σ_0, in the sense that certain distances in Assumption A.7(3) are bounded above by a sufficiently small ψ.

(4) (Smooth b with Lipschitz inverse) It is known that $b \in \mathcal{K}$ and

$$\sup_{x\in\mathbb{R}^J} \|(Db(x))^{-1}\|_{\text{op}} < \infty$$

where $Db(x)$ is the Jacobian of b.

The key assumption in Assumption 7 is that b is known to fall in a "well-behaved" function class $\mathcal{K}$, which—loosely speaking—consists of functions that deviate from a linear map by a suitably smooth function which has s derivatives and each derivative is in L^p. Because of this, we can limit Θ_I to just those functions h with conditional expectations in $\mathcal{K}$:

$$\Theta_I(\mathcal{K}) = \{h : \mathbb{E}[h(S,P) \mid X, Z] \in \mathcal{K}\} \ni a(\sigma^{-1}(\cdot,\cdot)).$$

The other regularity conditions in Assumption 7 establish that $\mathcal{K}$ is sufficiently small and the operator $u \mapsto \mathbb{E}[u(\delta) \mid X]$ is sufficiently well-behaved so that $\mathcal{K}$ can be entirely populated by conditional expectations of functions of δ.[19] That is, for any $k \in \mathcal{K}$, there exists some $H(\delta)$ such that $\mathbb{E}[H(\delta) \mid X] = k(X)$. Completeness would then imply a version of faithfulness with respect to $\mathcal{K}$. That is, for any candidate $H(\delta, P)$ where $\mathbb{E}[H(\delta, P) \mid X, Z] \in \mathcal{K}$, it follows that for some $H(\delta)$, we have

$$\mathbb{E}[H(\delta, P) \mid X, Z] = \mathbb{E}[H(\delta) \mid X, Z],$$

and therefore $H(\delta, P) = H(\delta)$ by completeness. We summarize this argument in the following result and verify that the conclusion of Proposition 1 continues to hold.

Proposition 6. *Under Assumptions 1 to 3 and 7, for any integrable $H(\delta, P)$, if $\mathbb{E}[H(\delta, P) \mid X, Z] \in \mathcal{K}$, then $H(\delta, P) = H(\delta)$ for some $H(\delta)$. As a result, price counterfactuals are identified.*

While the assumptions in Section 3.1 can be rationalized by economic restrictions on firm price-setting, Assumption 7 is more statistical in nature. We view Proposition 6 as an existence proof that restrictions on $\delta \mid X$ alone can prove faithfulness from completeness. Moreover, Assumption 7 is not knife-edge in nature, strongly suggesting that it can be further relaxed to allow for more flexible models for $\delta \mid X$. Overall, the combination of results in this section confirms that faithfulness can hold without either economic or statistical assumptions. We expect that many other sufficient conditions for faithfulness exist as well.

[19] Assumption 7(2) is satisfied by distributions with Gamma-like tails, though it rules out distributions with Gaussian0like tails. See Lemma A.7 and Remark A.3. We strongly suspect that these restrictions are not essential and can be further relaxed.

4. **Conclusion**

We have shown that price counterfactuals are identified without exogenous product characteristics in a general nonparametric demand model, given exogenous instruments that induce sufficient variation in prices. The richness of variation is captured by a new faithfulness condition, which essentially requires that the instruments make all causal price effects detectable. We show through a variety of non-nested sufficient conditions that faithfulness can hold without strong statistical or economic restrictions. We expect the new technical condition and results in this paper may also prove useful in other nonparametric models beyond demand.

We have not provided a theoretical analysis of nonparametric estimation based on these results, which would naturally require strengthening the identifying assumptions and incorporating additional regularity conditions (Compiani, 2022; Chen and Pouzo, 2015). As in Berry and Haile (2014), we leave developing this extension to future research. We nevertheless expect that our identification analysis will help guide empirical researchers towards more robust and credible estimation strategies, by clarifying the kinds of variation that can reveal counterfactuals in flexible demand models. Critically, in contrast to widely-held intuition since Berry and Haile (2014), our results suggest researchers can generally avoid conventional characteristic-based instruments—and potential biases associated with them—by looking for plausibly exogenous supply shocks and leveraging them via recentered instruments.

References

Ackerberg, Daniel A. and Gregory S. Crawford, "Estimating Price Elasticities in Differentiated Product Demand Models with Endogenous Characteristics," *Working Paper*, 2009.

Ai, Chunrong and Xiaohong Chen, "Efficient estimation of models with conditional moment restrictions containing unknown functions," *Econometrica*, 2003, *71* (6), 1795–1843.

Altonji, Joseph G. and Rosa L. Matzkin, "Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors," *Econometrica*, 2005, *73* (4), 1053–1102.

Andrews, Donald WK, "Examples of l 2-complete and boundedly-complete distributions," *Journal of Econometrics*, 2011, *199*, 213–220.

Andrews, Isaiah, Nano Barahona, Matthew Gentzkow, Ashesh Rambachan, and Jesse M Shapiro, "Structural estimation under misspecification: Theory and implications for practice," *The Quarterly Journal of Economics*, 2025, p. qjaf018.

Angrist, Joshua D., Kathryn Graddy, and Guido W. Imbens, "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish," *Review of Economic Studies*, 2000, *67* (3), 499–527.

Benkard, C Lanier and Steven Berry, "On the nonparametric identification of nonlinear simultaneous equations models: Comment on Brown (1983) and Roehrig (1988)," *Econometrica*, 2006, *74* (5), 1429–1440.

Berry, Steven, "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 1994, *25* (2), 242–262.

______ **, Amit Gandhi, and Philip Haile**, "Connected substitutes and invertibility of demand," *Econometrica*, 2013, *81* (5), 2087–2111.

______ **and Philip Haile**, "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, 2014, *82* (5), 1749–1797.

______ **and** ______ , "Nonparametric Identification of Differentiated Products Demand Using Micro Data," *Econometrica*, 2024, *92* (4), 1135–1162.

______ **, James Levinsohn, and Ariel Pakes**, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

______ **,** ______ **, and** ______ , "Voluntary export restraints on automobiles: Evaluating a trade policy," *American Economic Review*, 1999, *89* (3), 400–431.

Berry, Steven T and Philip A Haile, "Foundations of demand estimation," in "Handbook of industrial organization," Vol. 4, Elsevier, 2021, pp. 1–62.

Blundell, Richard, Joel L Horowitz, and Matthias Parey, "Measuring the price responsiveness of gasoline demand," *Quantitative Economics*, 2012, *3* (1), 29–51.

Borusyak, Kirill and Peter Hull, "Nonrandom exposure to exogenous shocks," *Econometrica*, 2023, *91* (6), 2155–2185.

______ **, Mauricio Caceres Bravo, and Peter Hull**, "Estimating demand with recentered instruments," *arXiv preprint arXiv:2504.04056*, 2025.

Brown, Donald J. and Rosa L. Matzkin, "Estimation of Nonparametric Functions of Simultaneous Equations Models, with an Application to Consumer Demand," *Cowles Foundation Discussion Papers 1175*, 1998.

Chen, Jiafeng, "Reinterpreting demand estimation," *arXiv preprint arXiv:2503.23524*, 2025.

Chen, Xiaohong and Demian Pouzo, "Sieve Wald and QLR inferences on semi/nonparametric conditional moment models," *Econometrica*, 2015, *83* (3), 1013–1079.

Chernozhukov, Victor, Guido W. Imbens, and Whitney K. Newey, "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics*, 2007, *139* (1), 4–14.

Chiappori, Pierre-André, Ivana Komunjer, and Dennis Kristensen, "Nonparametric identification and estimation of transformation models," *Journal of Econometrics*, 2015, *188* (1), 22–39.

Compiani, Giovanni, "Market counterfactuals and the specification of multiproduct demand: A nonparametric approach," *Quantitative Economics*, 2022, *13* (2), 545–591.

D'Haultfoeuille, Xavier, "On the completeness condition in nonparametric instrumental problems," *Econometric Theory*, 2011, *27* (3), 460–471.

D'Haultfœuille, Xavier and Philippe Février, "Identification of nonseparable triangular models with discrete instruments," *Econometrica*, 2015, *83* (3), 1199–1210.

Gandhi, Amit and Jean-François Houde, "Measuring Substitution Patterns in Differentiated-Products Industries," 2019. NBER Working Paper 26375.

Graddy, Kathryn, "Testing for imperfect competition at the Fulton fish market," *The RAND Journal of Economics*, 1995, pp. 75–92.

Grafakos, Loukas et al., *Classical fourier analysis*, Vol. 2, Springer, 2008.

Hörmander, Lars, *The analysis of linear partial differential operators I: Distribution Theory and Fourier Analysis*, Springer Berlin, Heidelberg, 2003.

Imbens, Guido W. and Whitney K. Newey, "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," *Econometrica*, 2009, *77* (5), 1481–1512.

Miao, Wang, Zhi Geng, and Eric J Tchetgen Tchetgen, "Identifying causal effects with proxy variables of an unmeasured confounder," *Biometrika*, 2018, *105* (4), 987–993.

Newey, Whitney K and James L Powell, "Instrumental variable estimation of nonparametric models," *Econometrica*, 2003, *71* (5), 1565–1578.

Newey, Whitney K., James L. Powell, and Francis Vella, "Nonparametric Estimation of Triangular Simultaneous Equations Models," *Econometrica*, 1999, *67* (3), 565–603.

Petrin, Amil, Mark Ponder, and Boyoung Seo, "Identification and estimation of discrete choice demand models when observed and unobserved characteristics are correlated," 2022. NBER Working Paper 30778.

Schumann, Aidan, "Multivariate bell polynomials and derivatives of composed functions," *arXiv preprint arXiv:1903.03899*, 2019.

Spirtes, Peter, Clark N Glymour, and Richard Scheines, *Causation, prediction, and search*, MIT press, 2000.

Torgovitsky, Alexander, "Identification of nonseparable models using instruments with small support," *Econometrica*, 2015, *83* (3), 1185–1197.

Wright, Philip Green, *The tariff on animal and vegetable oils* number 26, Macmillan, 1928.

Appendix for "Nonparametric Identification of Demand without Exogenous Product Characteristics"

Borusyak, Hull, Chen, and Lei

December 2025

Contents

Appendix A. Proofs for results in the main text

A.1. **Setup.** We first restate the assumptions and results in Section 2 more formally. Let F^* denote the distribution of (S, P, δ, X, Z). We assume for each F^* there is some σ such that $S = \sigma(\delta, P)$ almost-surely. Let F be the distribution of (S, P, X, Z). The assumptions will restrict the class of $\mathcal{F}^* \ni F^*$, which we keep implicit. Identification of price counterfactuals is formally defined in the following sense: Counterfactuals are identified if there exists some function of the observed data that predicts them, and these predictions are correct on a set of F^*-probability 1.

Definition 1. *We say that price counterfactuals are identified at F if there exists a mapping $C(s, p, p')$ such that for any F^* generating F, with a corresponding $\sigma = \sigma_{F^*}$,*

$$C(s, p, p') = \sigma(\sigma^{-1}(s, p), p')$$

for all values $(s, p, \delta, x, z), (s', p', \delta, x, z) \in E$ *for some set* E *with probability 1,* $\mathbb{P}_{F^*}(E) = 1$*. We say that price counterfactuals are identified if they are identified at all* F *generated by some* F^**.*

We first restate Assumption 1:

Assumption A.1. σ *is invertible in* δ*: For some measurable* σ^{-1}*,* $\delta = \sigma^{-1}(S, P)$ *almost surely. Assume that* $\mathbb{E}[\|\sigma^{-1}(S, P)\|] < \infty$*.*

We next restate Assumptions 2 and 3. Here, completeness is defined relative to all integrable functions. Weakening completeness by changing integrable to square-integrable does not affect subsequent results, so long as we modify them accordingly.

Assumption A.2. *For all* F*, the conditional distribution* $(S, P) \mid (X, Z)$ *is complete with respect to integrable functions: For any integrable* h*,*

$$\mathbb{E}_F[h(S, P) \mid X, Z] = 0 \implies h(S, P) = 0 \text{ almost surely.}$$

Assumption A.3. *For each* F^**,* $Z \perp\!\!\!\perp \delta \mid X$ *under* F^**.*

Fix some class of functions $\mathcal{H}$ mapping (S, P) to $\mathbb{R}^J$. Suppose $0 \in \mathcal{H}$ and all $h \in \mathcal{H}$ are integrable at all F.[1] Moreover, suppose all $h(s, p) \in \mathcal{H}$ are invertible in s: There exists h^{-1} such that, almost surely,

$$S = h^{-1}(h(S, P), P).$$

Fix some class of functions $\mathcal{K}$ mapping X to $\mathbb{R}^J$. Define an identified set for σ^{-1} with respect to $\mathcal{H}$ and $\mathcal{K}$:

$$\Theta_I = \Theta_I(\mathcal{H}, \mathcal{K}, F) = \{h \in \mathcal{H} : \mathbb{E}_F[h(S, P) \mid X, Z] \in \mathcal{K}\}.$$

We will examine throughout identification relative to these function classes $\mathcal{H}$ and $\mathcal{K}$, representing a priori restrictions (e.g., integrability, smoothness, etc.) that a researcher is willing to make.

A.2. **Suffices to identify σ^{-1} up to transformation.** The following formalizes Lemma 1: That ensuring Θ_I contains solely functions of the form $T(\sigma^{-1}(\cdot))$ suffices for identifying price counterfactuals.

Lemma A.1. *Assume that for every* F^**,*

(1) *Every* $h \in \Theta_I$ *is of the form* $h(S, P) = T(\sigma^{-1}(S, P))$ *almost surely for some measurable* $T : \mathbb{R}^J \to \mathbb{R}^J$*.*
(2) Θ_I *is nonempty.*

[1] We can also change this to square integrable throughout.

Then counterfactuals are identified by $C(s,p,p') = h^{-1}(h(s,p),p')$ *for any* $h \in \Theta_I$.

Proof. Fix any $h \in \Theta_I$. The assumptions imply that there exists a probability-one set E such that

$$s = h^{-1}(h(s,p),p) \quad h(s,p) = T(\sigma^{-1}(s,p)) \quad \delta = \sigma^{-1}(s,p) \quad s = \sigma(\delta,p)$$

for all $(s,p,\delta,x,z) \in E$. Therefore

$$s = h^{-1}\left(T(\delta),p\right).$$

Likewise, for some $(s',p',\delta,x,z) \in E$,

$$\sigma(\delta,p') = s' = h^{-1}(T(\delta),p') = h^{-1}(h(s,p),p'). \qquad \square$$

A.3. **Suffices to impose faithfulness.** Let $\mathcal{F}$ denote a class of functions mapping from (δ, P) to $\mathbb{R}^J$. We first restate Assumption 4 taking into account $\mathcal{F}$ and $\mathcal{K}$:

Assumption A.4. *Under Assumption A.3, the conditional distribution* $(\delta,P) \mid (X,Z)$ *satisfies faithfulness with respect to* $\mathcal{F}$ *and* $\mathcal{K}$*: For any* $H \in \mathcal{F}$ *such that* $\mathbb{E}_{F^*}[H(\delta,P) \mid X,Z] \in \mathcal{K}$*, we have that* $H(\delta,P) = H(\delta)$ *almost surely for some measurable* $H(\delta)$*.*

We now formalize Proposition 1—that faithfulness is sufficient for identification—relative to these definitions and assumptions:

Proposition A.1. *Fix* $(\mathcal{H},\mathcal{K},\mathcal{F})$*. Suppose Assumptions A.1, A.3, and A.4 hold for every* F^**. Assume further that for every* F^**,*

(1) There exists some invertible $T_0(\cdot)$ *such that* $\mathbb{E}_{F^*}[T_0(\delta) \mid X,Z] \in \mathcal{K}$ *and* $T_0(\sigma^{-1}(s,p)) \in \mathcal{H}$*.*

(2) For every $h \in \mathcal{H}$*,* $h(\sigma(\delta,p),p) \equiv H(\delta,p) \in \mathcal{F}$*.*

Then price counterfactuals are identified.

Proof. By Lemma A.1, we need to check assumptions (1)–(2) in Lemma A.1. First, condition (1) assumed implies that $T_0(\sigma^{-1}(s,p)) \in \mathcal{H}$ and its conditional expectation lies in $\mathcal{K}$. Thus $T_0(\sigma^{-1}(s,p)) \in \Theta_I$. Thus Θ_I is nonempty; this verifies (2) in Lemma A.1. Now, fix $h \in \Theta_I$, by the assumed condition (2), $h(\sigma(\delta,p),p) \in \mathcal{F}$. By Assumption A.4, we have that

$$h(S,P) = H(\delta) = H(\sigma^{-1}(S,P))$$

almost surely. This verifies assumption (1) in Lemma A.1. $\square$

A.4. **Faithfulness verification from discrete support.** We verify completeness implies faithfulness if X and δ are discrete with same-sized supports (Proposition 2).

Proposition A.2. *For every F^*, suppose the support of δ and X are both finite sets of size $M \in \mathbb{N}$. Let*

(1) $\mathcal{F}$ be all integrable functions of (δ, P),
(2) $\mathcal{H}$ be all integrable functions of (S, P) invertible in S
(3) $\mathcal{K}$ be all $\mathbb{R}^J$-valued functions of X.

Then, with these choices,

(1) Assumption A.4 follows from Assumption A.2 and Assumption A.3.
(2) The assumptions of Proposition A.1 are satisfied.

Proof. Note that any real-valued function $h(\delta)$ and $g(x)$ can be represented as a vector in $\mathbb{R}^M$. Consider

$$\mathbb{E}[h(\delta) \mid X, Z] = \mathbb{E}[h(\delta) \mid X] \equiv g(X).$$

If we represent both as vectors in $\mathbb{R}^M$, then we have

$$Q'_{\delta|X} h = g,$$

for a matrix $Q_{\delta|X}$ whose $(i,j)^{\text{th}}$ entry is $\mathbb{P}_{F^*}(\delta = \delta_i \mid X = x_j)$, where we number values in the support of δ as δ_i and likewise for values of X. The matrix $Q_{\delta|X}$ is square by assumption. Completeness implies that $Q'_{\delta|X}$ is full-rank, and thus invertible.

Now, consider any function where $\mathbb{E}[H(\delta, P) \mid X, Z] = k(X) \in \mathcal{K}$. For each coordinate j, there is a function $H_{0j}(\delta)$ which can be represented as $H_{0j} = (Q'_{\delta|X})^{-1} k_j$. Construct $H_0 = (H_{01}, \dots, H_{0J})'$. By construction $\mathbb{E}[H_0(\delta) \mid X, Z] = k(X)$. Completeness implies that $H(\delta, P) = H_0(\delta)$. This shows (1).

For (2), we can easily check that $\sigma^{-1} \in \mathcal{H}$ and $\mathbb{E}[\sigma^{-1} \mid X, Z] = \mathbb{E}[\delta \mid X] \in \mathcal{K}$. Thus (1) in Proposition A.1 is satisfied with $T_0(\delta) = \delta$. If $h(S, P) \in \mathcal{H}$ is integrable, then so is $H(\delta, P) = h(\sigma(\delta, P), P)$ since $S = \sigma(\delta, P)$ a.s. Thus $H \in \mathcal{F}$. This verifies (2) in Proposition A.1. ☐

A.5. **Faithfulness verification from λ index.** Next, we state and verify the case (Proposition 4) where

$$P \perp\!\!\!\perp (X, Z) \mid \lambda(X, Z), \delta.$$

We omit Proposition 3 since it is a corollary of Proposition 4 with $P = \lambda(X, Z) = Z$.

Assumption A.5. *Let $\lambda : \mathbb{R}^J \times \mathcal{Z} \to \mathcal{L} \subset \mathbb{R}^d$ be an index that is invertible in Z.*

(1) For such a λ, $P \perp\!\!\!\perp (X, Z) \mid \lambda(X, Z), \delta$.

(2) Let $\lambda = \lambda(X, Z)$. The joint distribution of (X, λ) has a density with respect to some product measure $\mu_X \otimes \mu_\lambda$ over $\mathcal{X} \times \mathcal{L}$, which is strictly positive for $\mu_X \otimes \mu_\lambda$-almost every $(x, \lambda) \in \mathcal{X} \times S_\lambda$ for some $S_\lambda \subset \mathcal{L}$ and $(\mu_X \otimes \mu_\lambda)(\mathbb{R}^J \times S_\lambda) > 0$.

Proposition A.4. *Let*

(1) $\mathcal{F}$ be all integrable functions of (δ, P)
(2) $\mathcal{H}$ be all integrable functions of (S, P) invertible in S
(3) $\mathcal{K}$ be all integrable functions of X

Then,

(1) Assumption A.4 follows from Assumption A.2 and Assumption A.3.
(2) The assumptions of Proposition A.1 are satisfied.

Proof. (1) Take $H \in \mathcal{F}$ with

$$\mathbb{E}\left[H(\delta, P) \mid X, Z\right] = k(X) \in \mathcal{K}.$$

By law of iterated expectations

$$k(X) = \mathbb{E}[\underbrace{\mathbb{E}[H(\delta, P) \mid \delta, \lambda]}_{\overline{H}(\delta,\lambda)} \mid X, Z] = \mathbb{E}[\overline{H}(\delta, \lambda) \mid X, \lambda]$$

since (X, Z) and (X, λ) are measurable with respect to each other.

By the assumptions, there exists a set $Q \subset \mathcal{X} \times S_\lambda$ such that for every $(x, \lambda) \in Q$, $k(x) = g(x, \lambda)$ where

$$g(x, \lambda) = \mathbb{E}[\overline{H}(\delta, \lambda) \mid X = x]. \qquad (\delta \perp\!\!\!\perp \lambda \mid X)$$

This set Q can be taken such that

$$(\mu_X \otimes \mu_\lambda)\left\{Q^C \cap (\mathcal{X} \times S_\lambda)\right\} = 0. \tag{A.1}$$

By (A.1), we have that for μ_λ-almost every $\lambda \in S_\lambda$,

$$g(X, \lambda) = \mathbb{E}[\overline{H}(\delta, \lambda) \mid X] = k(X) \quad \mu_X\text{-almost everywhere.}$$

As a result, there exists some λ_0 such that

$$g(X, \lambda_0) = \mathbb{E}[\overline{H}(\delta, \lambda_0) \mid X] = k(X) \quad F^*\text{-a.s.}$$

Since $k(X)$ is integrable, so is $\overline{H}(\delta, \lambda_0)$. Finally,

$$0 = \mathbb{E}[\overline{H}(\delta, \lambda_0) - H(\delta, P) \mid X, Z].$$

By completeness, we have that $\overline{H}(\delta, \lambda_0) = H(\delta, P)$ a.s. This verifies Assumption A.4.

(2) The proof is analogous to Proposition A.2(2). ∎

A.6. **Faithfulness verification from price separability.** We next restate Assumption 6 and Proposition 5, which rely on restrictions about derivatives of P in z. (A.2) is slightly different from—but is implied by—its counterpart in Assumption 6.

Assumption A.6. *We have the following:*

(1) *(Support) For almost every δ, there is a connected open set $U_\delta \subset \mathbb{R}^J$ such that $P \mid \delta$ has a density $f(p \mid \delta) > 0$ on U_δ and $\mathbb{P}(P \in U_\delta \mid \delta) = 1$. The distribution $(S, P, Z) \mid X$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^{J\times J\times d_z}$. There exists an open set $U \subset \mathbb{R}^{J\times J\times d_z}$ where the density of $(S, P, Z) \mid X$ is positive on U and $P((S, P, Z) \in U \mid X) = 1$ a.s.*

(2) *(Price equation) The random variable P can be represented as*

$$P = f(X, \delta, Z, \omega) \text{ for } (\omega, \delta) \perp\!\!\!\perp Z \mid X \tag{A.2}$$

for some function f continuously differentiable in Z with Jacobian $D_z f$. The joint distribution $(\omega, \delta) \mid X$ has density $f(\omega, \delta \mid x)$ with respect to some dominating measure λ_x.

(3) *(Price separability) The Jacobian of f in Z satisfies the following: For measurable functions A, B,*

$$\underbrace{D_z f(X, \delta, Z, \omega)}_{J\times d_z} = \underbrace{A(f(X, \delta, Z, \omega), \delta)}_{J\times J} \cdot \underbrace{B(X, Z)}_{J\times d_z} \tag{A.3}$$

where $A(f(X, \delta, Z, \omega), \delta)$ and $B(X, Z)$ are both full-rank a.s. with $d_z \geq J$

(4) *(Domination) There exists $G_1(x, \delta, \omega) \geq 0$ and $G_2(x, \delta, \omega) \geq 0$ such that a.s.,*[2]

$$\|\sigma_p(\delta, P) f_z(X, Z, \delta, \omega)\| \leq G_1(X, \delta, \omega)$$
$$\|f_z(X, Z, \delta, \omega)\| \leq G_2(X, \delta, \omega)$$

and $\mathbb{E}[G_1^2 + G_2^2 \mid X] = \mathbb{E}[G_1^2 + G_2^2 \mid X, Z] < \infty$.

Proposition A.5. *Let*

(1) *$\mathcal{F}$ denote the class of functions $H(\delta, P)$ where (i) $\mathbb{E}[H^2 \mid X, Z] < \infty$, (ii) for F^*-almost every δ, $H(\delta, \cdot)$ is continuously differentiable with derivative $H_p(\delta, \cdot)$ on U_δ, and (iii)*

$$\frac{\partial}{\partial z}\mathbb{E}[H(\delta, f(X, \delta, z, \omega)) \mid X, Z = z] = \mathbb{E}[H_p(\delta, f(X, \delta, z, \omega)) D_z f(X, \delta, z, \omega) \mid X].$$

(Interchange of differentiation and expectation)

[2]For concreteness, we choose the Frobenius norm as the matrix norm.

(2) $\mathcal{H} \subset L^2_F(S, P)$ be the set of $\mathbb{R}^J$-valued functions $h(s, p)$ such that (i) h is continuously differentiable in (s, p) on U, (ii) the derivatives are locally Lipschitz: There exists $\epsilon > 0$ such that for each $j = 1, \ldots, J$ and all $\|(s, p) - (S, P)\| < \epsilon$,

$$\max\left(\left\|\frac{\partial h_j}{\partial s}(s, p) - \frac{\partial h_j}{\partial s}(S, P)\right\|, \left\|\frac{\partial h_j}{\partial p}(s, p) - \frac{\partial h_j}{\partial p}(S, P)\right\|\right) \leq L(S, P)\|(s, p) - (S, P)\|$$

where $\mathbb{E}[L^2(S, P) \mid X, Z] < \infty$, and (iii) The derivatives h_s, h_p are integrable:

$$\mathbb{E}[\|h_s(S, P)\|^2 + \|h_p(S, P)\|^2 \mid X, Z] < \infty.$$

(iv) h is invertible in S.

(3) $\mathcal{K}$ be all square-integrable functions of X.

Assume $\sigma^{-1} \in \mathcal{H}$. Suppose Assumptions A.1 to A.3 and A.6 hold. Then,

(1) Assumption A.4 holds

(2) The assumptions in Proposition A.1 hold

Proof. (1) Let $H \in \mathcal{F}$. Write

$$k(x) = \mathbb{E}[H(\delta, P) \mid X = x, Z = z] = \int H(\delta, f(x, \delta, z, \omega))\, f(\omega, \delta \mid x) d\lambda_x.$$

Since $\mathcal{H} \in \mathcal{F}$, differentiating both sides in z yields

$$0 = \mathbb{E}[H_p(\delta, P) A(P, \delta) \mid X, Z] \cdot B(X, Z).$$

Since $d_z \geq J$ and $B(X, Z)$ is full rank, we have that almost surely

$$\mathbb{E}[H_p(\delta, P) A(P, \delta) \mid X, Z] = 0.$$

By completeness,

$$H_p(\delta, P) A(P, \delta) = 0 \implies H_p(\delta, P) = 0 \text{ a.s.}$$

since A is full rank. Since H_p is continuous in p and P is supported with positive density on U_δ, $H_p(\delta, p) = 0$ on U_δ. As a result, $H(\delta, \cdot)$ is constant on U_δ for almost every δ. Hence there exist some $H(\delta)$ where $H(\delta) = H(\delta, P)$ almost surely. This verifies Assumption A.4.

(2) Lemma A.2 verifies that Proposition A.1(2) holds. Finally, Proposition A.1(1) holds by choosing $T_0(d) = d$, since

$$\sigma^{-1} \in \mathcal{H}, \quad \mathbb{E}[\sigma^{-1} \mid X, Z] = k(X) \in L^2_F(X) = \mathcal{K},$$

under the assumption that σ^{-1} is square-integrable.

□

Lemma A.2. *Suppose Assumption A.6 holds. Let $\mathcal{H}, \mathcal{F}$ be defined as in Proposition A.5 and suppose $\sigma^{-1} \in \mathcal{H}$. Define*

$$H(\delta, p) = h(\sigma(\delta, p), p)$$

Then $H \in \mathcal{F}$.

Proof. It suffices to show that differentiation in z and expectation are interchangeable for H, since the other properties of $\mathcal{F}$ are assumed in $\mathcal{H}$. Since it suffices to show the claim entrywise over entries of h, we abuse notation and denote $h(s, p)$ as an entry and treat it as a scalar function.

We let f_z denote $D_z f$. Define $h^*(\delta, x, z, \omega) = h(\sigma(\delta, f(x, \delta, z, \omega)), f(x, \delta, z, \omega))$. Fix some z_0. Let

$$h_z^*(\delta, X, Z, \omega) = (h_s(S, P)'\sigma_p(\delta, P) + h_p(S, P)')\, f_z(X, \delta, Z, \omega)$$

be the (transposed) gradient at (δ, X, Z, ω). We show that if $h \in \mathcal{H}$ then

$$\frac{\partial}{\partial z}\mathbb{E}[h^*(\delta, X, Z, \omega) \mid X, Z] = \mathbb{E}[h_z^*(\delta, X, Z, \omega) \mid X, Z].$$

It suffices to show that

$$\lim_{t\to 0}\mathbb{E}\left[\sup_{\|v\|=1}\left|\frac{h^*(\delta, X, z_0 + tv, \omega) - h^*(\delta, X, z_0, \omega)}{t} - h_z^*(\delta, X, z, \omega)v\right| \mid X\right] = 0. \quad \text{(A.4)}$$

(A.4) implies

$$\lim_{t\to 0}\sup_{\|v\|=1}\left|\mathbb{E}\left[\frac{h^*(\delta, X, z_0 + tv, \omega) - h^*(\delta, X, z_0, \omega)}{t} \mid X\right] - \mathbb{E}[h_z^*(\delta, X, z, \omega) \mid X]v\right| = 0,$$

which implies that $z \mapsto \mathbb{E}[h^*(\delta, X, z, \omega) \mid X]$ is (Frechet) differentiable at z_0 and its gradient is equal to $\mathbb{E}[h_z^*(\delta, X, z, \omega) \mid X]$.

Towards (A.4), at a fixed (δ, x, ω), observe that since h^* is differentiable at z_0, we have pointwise convergence

$$\lim_{t\to 0}\sup_{\|v\|=1}\left|\frac{h^*(\delta, x, z_0 + tv, \omega) - h^*(\delta, x, z_0, \omega)}{t} - h_z^*(\delta, x, z_0, \omega)v\right| = 0.$$

Thus it suffices to show the following and apply the dominated convergence theorem: For all sufficiently small t,

$$\sup_{\|v\|=1}\left|\frac{h^*(\delta, X, z_0 + tv, \omega) - h^*(\delta, X, z_0, \omega)}{t} - h_z^*(\delta, X, z_0, \omega)v\right| \le G_0(\delta, X, \omega) \quad \text{(A.5)}$$

where $\mathbb{E}[G_0(\delta, X, \omega) \mid X] < \infty$.

Observe that

$$\|h_z^*(\delta, x, z_0, \omega)\| \leq \|h_s(s,p)\|\|\sigma_p(\delta,p)f_z(x,z_0,\delta,\omega)\| + \|h_p(s,p)\|\|f_z(x,z_0,\delta,\omega)\|$$
$$\leq \|h_s(s,p)\|G_1(x,\delta,\omega) + \|h_p(s,p)\|G_2(x,\delta,\omega)$$

for $s = \sigma(\delta, x, z_0, \omega), p = f(x, \delta, z_0, \omega)$. Since the derivatives of h and G_1, G_2 are all square-integrable, by Cauchy-Schwarz $\|h_s(s,p)\|G_1(x,\delta,\omega) + \|h_p(s,p)\|G_2(x,\delta,\omega)$ is an integrable dominating function.

Therefore it suffices to show that the difference quotient can be dominated

$$\sup_{\|v\|=1} \left| \frac{h^*(\delta, X, z_0 + tv, \omega) - h^*(\delta, X, z_0, \omega)}{t} \right| \leq G_0(\delta, X, \omega) \tag{A.6}$$

Towards (A.6), define $p_{tv} = f(\delta, x, z_0 + tv, \omega)$ and $s_{tv} = \sigma(\delta, p_{tv})$. By the mean-value theorem in h, we have

$$h^*(\delta, x, z_0+tv, \omega) - h^*(\delta, X, z_0, \omega) = h_s(\tilde{s}_{t,v}, \tilde{p}_{t,v})(s_{tv} - s_0) + h_p(\tilde{s}_{t,v}, \tilde{p}_{t,v})(p_{tv} - p_0) \tag{A.7}$$

where $(\tilde{s}_{t,v}, \tilde{p}_{t,v})$ is some point on the line segment connecting (s_0, p_0) with (s_{tv}, p_{tv}).

We can write $h_s(\tilde{s}_{tv}, \tilde{p}_{tv}) = h_s(s_0, p_0) + R_s(s_0, p_0)$ where $\|R_s(s_0,p_0)\| \leq L(s_0,p_0)t$ for all $t < \epsilon$. Similarly for h_p. Thus

$$(A.7) = h_s(s_0,p_0)(s_{tv} - s_0) + R_s(s_0,p_0)(s_{tv} - s_0) + h_p(s_0,p_0)(p_{tv} - p_0) + R_p(s_0,p_0)(p_{tv} - p_0)$$

By the mean-value theorem applied to $z \mapsto \sigma(\delta, f(\delta, x, z, \omega))$, we have that

$$s_{tv} - s_0 = \{\sigma_p(\delta, \check{p}_{tv}) f_z(x, \check{z}, \delta, \omega)\}\, tv$$

where $\check{z}$ is on the line segment between $z_0, z_0 + tv$ and $\check{p}_{tv} = f(x, \check{z}, \delta, \omega)$. We thus have that

$$\|s_{tv} - s_0\| \leq tG_1(x, \delta, \omega).$$

Similarly, we have that

$$\|p_{tv} - p_0\| \leq tG_2(x, \delta, \omega).$$

Thus, for all $t < \epsilon$,

$$\sup_{\|v\|=1} \left\| \frac{(A.7)}{t} \right\| \leq \|h_s(s_0,p_0)\|G_1(x,\delta,\omega) + \|h_p(s_0,p_0)\|G_2(x,\delta,\omega)$$
$$+ tL(s_0,p_0)\{G_1(x,\delta,\omega) + G_2(x,\delta,\omega)\}$$

The right-hand side is a function of (ω, δ, x) that is integrable by Cauchy–Schwarz. This concludes the proof. □

Lemma A.3. *Let $(\omega, \delta) \perp\!\!\!\perp Z \mid X$ and suppose P is generated by (12). Assume the dimension of z is at least J. Assume that*

(1) *The codomains of f_0, f_1, f_2 are all $\mathbb{R}^J$.*
(2) *$f_0(\cdot, \delta)$ is invertible where $f_0^{-1}(\cdot, \delta)$ is continuously differentiable for every δ with invertible Jacobian $Q_0(p, \delta) \in \mathbb{R}^{J \times J}$.*
(3) *f_1 is continuously differentiable in z with nonsingular Jacobian $Q_1(z, x) \in \mathbb{R}^{J \times d_z}$*

Then items 2 and 3 of Assumption A.6 are satisfied.

Proof. Write

$$f_0^{-1}(p, \delta) = f_1(z, x) + f_2(\delta, x, w).$$

Differentiate in z implicitly to obtain

$$Q_0(p, \delta) D_z f = Q_1(z, x).$$

Since Q_0 is assumed to be invertible, we have

$$D_z f = \underbrace{Q_0^{-1}(p, \delta)}_{A} \underbrace{Q_2(z, x)}_{B}.$$

By assumption A and B are not singular. □

A.7. **Faithfulness verification from location-scale model for δ.**

A.7.1. *Mathematical preliminaries.* Let $\mathbb{N}$ denote the set of nonnegative integers and $\mathbb{S}_+^J$ denote the set of $J \times J$ positive semidefinite matrices. For any multi-index $\alpha = (\alpha_1, \ldots, \alpha_J) \in \mathbb{N}^J$ and function $f(x) : \mathbb{R}^J \mapsto \mathbb{R}^s$, where $x \in \mathbb{R}^J$ and s may be different from J, we write $|\alpha| = \sum_{j=1}^J \alpha_j$, $\alpha! = \prod_{j=1}^J \alpha_j!$, and

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_J^{\alpha_J}, \quad D_x^\alpha f(x) = D_{x_1}^{\alpha_1} D_{x_2}^{\alpha_2} \cdots D_{x_J}^{\alpha_J} f(x).$$

Furthermore, for any function $g : \mathbb{R}^J \mapsto \mathbb{R}$, we define its Fourier transform as

$$\hat{g}(\omega) = \int_{\mathbb{R}^J} e^{-i\langle \omega, u \rangle} g(u) du.$$

For a function class $\mathcal{L}$, $f \in \mathcal{L}^{\otimes J}$ iff $f = (f_1, \ldots, f_J)'$ where $f_j \in \mathcal{L}$.

Given a function $v : \mathbb{R}^J \mapsto \mathbb{R}$ for which v is positive everywhere, let $\mathbb{W}^{s,p}(\mathbb{R}^J)$ denote the Bessel potential space of indices (s, p):

$$\mathbb{W}^{s,p}(\mathbb{R}^J; \mathbb{R}^J) = \left\{ u : \mathbb{R}^J \to \mathbb{R}^J : (1 + \|\omega\|_2^2)^{s/2} \hat{u}_j(\omega) \in L^p(\mathbb{R}^J),\ j \in [J] \right\}. \quad \text{(A.8)}$$

For any $s = m + \sigma$ where $m = \lfloor s \rfloor$ is the integer part and $\sigma \in [0, 1)$ is the fractional part, let $\|\cdot\|_{\mathbb{W}^{s,p}}$ denote the associated norm with

$$\|u\|_{\mathbb{W}^{s,p}} = \sum_{|\alpha| \leq m} \|D^\alpha u\|_{L^p} + \sum_{\alpha : |\alpha| = m} [D^\alpha u]_{W^{\sigma,p}},$$

where

$$[h]_{\mathbb{W}^{\sigma,p}} := \left(\int_{\mathbb{R}^J} \int_{\mathbb{R}^J} \frac{\|h(x) - h(y)\|^p}{\|x - y\|^{\sigma p + J}} dx dy \right)^{1/p}, \quad [h]_{\mathbb{W}^{\sigma,\infty}} := \sup_{x \neq y} \frac{\|h(x) - h(y)\|}{\|x - y\|^\sigma}.$$

Define the following space:

$$\begin{aligned} \mathbb{K}^s(\mathbb{R}^J; \mathbb{R}^J) &\equiv \text{lin}(\mathbb{R}^J; \mathbb{R}^J) \oplus \mathbb{W}^{s,\infty}(\mathbb{R}^J; \mathbb{R}^J) \\ &\equiv \{u : \mathbb{R}^J \to \mathbb{R}^J : u(x) = Ax + r(x), A \in \mathbb{R}^{J \times J}, r \in \mathbb{W}^{s,\infty}(\mathbb{R}^J, \mathbb{R}^J)\} \end{aligned} \tag{A.9}$$

where $\text{lin}(\mathbb{R}^J; \mathbb{R}^J)$ is the space of all bounded linear functions that map $\mathbb{R}^J$ to $\mathbb{R}^J$. In the following, for notational convenience, we suppress the spaces $\mathbb{R}^J$ or $\mathbb{R}^J; \mathbb{R}^J$ in $\mathbb{W}^{s,\infty}$ and $\mathbb{K}^s$. For any $s_0 \leq s_1$, since $\mathbb{W}^{s_1,\infty} \subset \mathbb{W}^{s_0,\infty}$, we have $\mathbb{K}^{s_1} \subset \mathbb{K}^{s_0}$.

Last, for a linear operator $\mathcal{T}$ that maps a Banach space $\mathcal{U}$ to a Banach space $\mathcal{V}$, let

$$\|\mathcal{T}\|_{\mathcal{U} \mapsto \mathcal{V}} = \sup_{u \in \mathcal{U}} \frac{\|\mathcal{T} u\|_{\mathcal{V}}}{\|u\|_{\mathcal{U}}}.$$

A.7.2. *Assumptions.*

Assumption A.7. *There exists $s, M \in \mathbb{N}$ with $M = J + 2s + 1$ and $s > J + 1$ such that the following holds:*

(1) There exist invertible $a, b : \mathbb{R}^J \to \mathbb{R}^J$ with

$$a(\delta) = b(X) + \Sigma(X)\xi \tag{A.10}$$

where $\xi \perp\!\!\!\perp X$ and $\Sigma : \mathbb{R}^J \mapsto \mathbb{S}^J_+$.

(2) Let q denote the density function of $-\xi$ with respect to the Lebesgue measure, which we assume to exist. There exist $0 < C_- < C_+ < \infty$ such that, for any $\upsilon \in \mathbb{R}^J$,

$$C_-(1 + \|\upsilon\|_2^2)^{-s/2} \leq |\hat{q}(\upsilon)| \leq C_+(1 + \|\upsilon\|_2^2)^{-s/2},$$

and, for any $\alpha \in \mathbb{N}^J$ with $|\alpha| \leq M$,

$$\|D^\alpha_\upsilon \hat{q}(\upsilon)\| \leq C_+(1 + \|\upsilon\|_2^2)^{-s/2 - |\alpha|/2}$$

(3) *There exist* $\Sigma_0 \in \mathbb{S}^J_+$, *and* $\psi \in (0, \lambda_{\min}(\Sigma_0)/2)$ *such that,*

$$\sup_{x \in \mathbb{R}^J} \|\Sigma(x) - \Sigma_0\|_{\text{op}} + \int_{\mathbb{R}^J} \|\Sigma(x) - \Sigma_0\|_{\text{op}} dx \leq \psi,$$

and, for any $\gamma \in \mathbb{N}^J$ *with* $|\gamma| \leq M$,

$$\sup_{x \in \mathbb{R}^J} \|D^\gamma_x \Sigma(x)\|_{\text{op}} + \int_{\mathbb{R}^J} \|D^\gamma_x \Sigma(x)\|_{\text{op}} dx \leq \psi.$$

(4) $b \in \mathbb{K}^M(\mathbb{R}^J)$ *and* $\sup_{x \in \mathbb{R}^J} \|(Db(x))^{-1}\|_{\text{op}} < \infty$.

Intuitively,

- Assumption A.7(1) assumes that $\delta \mid X$ follows a location-scale model.
- Assumption A.7(2) assumes that this familly has shape corresponding to some density $q(\cdot)$, and $|\hat{q}(v)| \asymp \|v\|^{-s}$ for large v in signal space, with derivatives up to M of the Fourier transform satisfying $\|D^\alpha_v \hat{q}(v)\| \lesssim \|v\|^{-s-|\alpha|}$.
- Assumption A.7(3) imposes that $\Sigma(x)$ is close in operator norm to some constant Σ_0 and similarly for the derivatives of $\Sigma(x)$.
- Assumption A.7(4) imposes that $b \in \mathbb{K}^M$ and its inverse b^{-1} is Lipschitz.

A.7.3. *Faithfulness verification.*

Proposition A.6. *Assume that Assumption A.7 holds with sufficiently small* ψ (A.11). *Let* M, s *be as in Assumption A.7. Let*

(1) $\mathcal{F}$ *be all integrable functions of* (δ, P)
(2) $\mathcal{H}$ *be all integrable functions of* (S, P) *invertible in* S
(3) $\mathcal{K} = \mathbb{K}^{2s}$

Then, with these choices,

(1) *Assumption A.4 follows from Assumptions A.2 and A.3*
(2) *The assumptions of Proposition A.1 are satisfied.*

Proof. Faithfulness (1) follows immediately from Theorem A.1 and Assumption A.2, applied entrywise. For Proposition A.1(1), observe that $b(X) = \mathbb{E}[a(\delta) \mid X, Z] \in \mathbb{K}^M \subset \mathbb{K}^{2s} = \mathcal{K}$ and $\sigma^{-1} \in \mathcal{H}$ with the choice that $T_0(\delta) = a(\delta)$. Likewise, Proposition A.1(2) holds immediately. □

We now outline the remaining argument:

(1) Theorem A.1 is the key result. Its proof argues that it is without loss to work with $\tilde{\delta} = \Sigma_0^{-1}\delta$, $\tilde{X} = \Sigma_0^{-1} b(X)$ such that $\tilde{\delta} = \tilde{X} + \tilde{\Sigma}(\tilde{X})\xi$ and $\tilde{\Sigma}(\tilde{X})$ is centered

around I. Moreover, it suffices to show that

$$\mathbb{E}[u(\delta) \mid X] = k(X) \quad k \in \mathbb{W}^{2s,\infty}$$

is solvable by some u.

With these normalizations, Theorem A.2 shows that there exists some $u(\delta)$ such that $\mathbb{E}[u(\delta) \mid X] = k(X)$, for scalar-valued functions (u, k). The bulk of the argument justifies Theorem A.2.

(2) Suppose $\Sigma(x) = I$. Then $\mathbb{E}[u(\delta) \mid X] = (u \star q)(x)$ is a convolution. Thus we can take u such that its Fourier transform is a ratio $\hat{u} = \frac{\hat{k}}{\hat{q}}$. Lemma A.4 verifies that u is a proper function.

(3) Next, with $\Sigma(x) \neq I$, one could consider the deviation of the conditional expectation operator from the $\Sigma = I$ case:

$$(\mathcal{T}u)(x) = (\mathcal{T}_0 u)(x) + (\mathcal{E}u)(x) = \left\{\mathcal{T}_0 \left(\mathrm{Id} + \mathcal{T}_0^{-1}\mathcal{E}\right) u\right\}(x)$$

for $\mathcal{T}$ the conditional expectation operator $\delta \mid X$ and $\mathcal{T}_0$ the conditional expectation operator under $\Sigma(x) = I$. Now, we have the geometric expansion

$$(\mathrm{Id} + \mathcal{T}_0^{-1}\mathcal{E})^{-1} = \sum_{\ell=0}^{\infty} \left(-\mathcal{T}_0^{-1}\mathcal{E}\right)^{\ell}$$

upon verifying that the right-hand side converges. Lemma A.5 derives the key condition for convergence. Given a k, one can then construct u as

$$u = \left(\sum_{\ell=0}^{\infty} \left(-\mathcal{T}_0^{-1}\mathcal{E}\right)^{\ell}\right) \mathcal{T}_0^{-1} k.$$

(4) The key condition for convergence turns out to be a bound on the discrepancy in Fourier space of the integration kernels: $\hat{q}(\Sigma(x)\omega) - \hat{q}(\omega)$. The proof of Theorem A.2 verifies that so long as $\Sigma(x)$ is sufficiently close to I under Assumption A.7(3), this bound is attainable.

Theorem A.1. *Assume that Assumption A.7 holds with*

$$\psi < \psi^* \equiv \psi^* \left(J, s, C_{\pm}, \lambda_{\min}(\Sigma_0), \|Db\|_{\mathbb{W}^{M-1,\infty}}, \|Db^{-1}\|_{L^\infty}\right) \tag{A.11}$$

for some constant ψ^* that only depends on the parameters inside the parentheses. Then for any $k \in \mathbb{K}^{2s}(\mathbb{R}^J, \mathbb{R}^J)$, there exists a proper function $u : \mathbb{R}^J \mapsto \mathbb{R}^J$ such that $\mathbb{E}[u(\delta) \mid X] = k(X)$ almost surely.

Proof. We first note that it suffices to prove the claim for scalar-valued $k \in \mathbb{K}^{2s}(\mathbb{R}^J;\mathbb{R})$ since we can concatenate the solutions. We argue that it is sufficient to prove Theorem A.2. The reason is that we can change variables and write $\tilde{X}$ for $b(X)$, $\tilde{\xi}$ for $\Sigma_0\xi$, and $\tilde{\Sigma}(\tilde{X}) = (\Sigma \circ b^{-1}(\tilde{X}))\Sigma_0^{-1}$.

We need to show that, for sufficiently small choices of ψ, $\tilde{\Sigma}(\cdot) - I$ is suitably bounded by some $\tilde{\psi}$ so that Assumption A.7(3) is satisfied for Theorem A.2. Lemma A.10 bounds $\tilde{\Sigma}(\cdot) - I$ in terms of $\Sigma(\cdot) - \Sigma_0$, where the bound depends only on the additional quantities $\|Db\|_{\mathbb{W}^{M-1,\infty}}, \|Db^{-1}\|_{L^\infty}, \lambda_{\min}^{-1}(\Sigma_0)$. Thus, there exists

$$\psi^* := \psi^*\left(J, s, C_\pm, \lambda_{\min}(\Sigma_0), \|Db\|_{\mathbb{W}^{M-1},\infty}, \|Db^{-1}\|_{L^\infty}\right)$$

such that, if $\psi \leq \psi^*$, we have that

$$\|\tilde{\Sigma}(\cdot) - I\|_{\mathbb{W}^{M,\infty}} + \|\tilde{\Sigma}(\cdot) - I\|_{\mathbb{W}^{M,1}} \leq \tilde{\psi}^*(J, s, C_\pm).$$

required by Theorem A.2.

Now, having checked the conditions, we can apply Theorem A.2 to $\tilde{X}$. Given $k \in \mathbb{K}^{2s}$, since b is invertible, we can write

$$k(X) = (k \circ b^{-1} \circ b)(X) = (k \circ b^{-1})(\tilde{X}).$$

By Lemma A.9, $k \circ b^{-1} \in \mathbb{K}^{2s}$. By definition,

$$(k \circ b^{-1})(\tilde{x}) = A\tilde{x} + \tilde{k}(\tilde{x}) \quad \tilde{k}(\tilde{x}) \in \mathbb{W}^{2s,\infty}.$$

By Theorem A.2, we can then find $\tilde{u}(\delta)$ for which

$$\mathbb{E}[\tilde{u}(\delta) \mid \tilde{X}] = \tilde{k}(\tilde{X}).$$

Let $u(\delta) = A\delta + \tilde{u}(\delta)$, we then have

$$\mathbb{E}[u(\delta) \mid \tilde{X}] = \mathbb{E}[u(\delta) \mid X] = (k \circ b^{-1})(\tilde{X}) = k(X).$$

□

Theorem A.2. *Assume that Assumption A.7 holds with $b(x) = x, \Sigma_0 = I$ and*

$$\psi < \psi^* \equiv \psi^*(J, s, C_\pm) \tag{A.12}$$

for some constant ψ^* that only depends on the parameters inside the parentheses, defined in (A.13). Then for any $k \in \mathbb{W}^{2s,\infty}$, there exists a proper function u such that $\mathbb{E}[u(\delta) \mid X] = k(X)$ almost surely.

Proof. By Lemmas A.4 and A.5, it remains to prove (A.17).

Step 1: Bound L^1 norm of $\Delta_\omega(x)$ and its derivatives by $(1+\|\omega\|^2)^{-s/2}$.

By the Taylor expansion,

$$\Delta_\omega(x) = \langle \nabla \hat{q}((\lambda\Sigma(x) + (1-\lambda)I)\omega), (\Sigma(x) - I)\omega\rangle$$

for some $\lambda \in (0,1)$ that depends on x. By Assumption A.7(3), $\psi < 1/2$ and

$$\|(\lambda\Sigma(x) + (1-\lambda)I)\omega\|_2 \geq (1-\psi)\|\omega\|_2 \geq \|\omega\|_2/2.$$

By Assumption A.7(2),

$$\begin{aligned}\|\nabla \hat{q}((\lambda\Sigma(x) + (1-\lambda)I)\omega)\|_2 &\leq C_+(1+\|\omega\|_2^2/4)^{-s/2-1/2} \\ &\leq C_+ 2^{s+1}(1+\|\omega\|_2^2)^{-s/2-1/2}.\end{aligned}$$

Then

$$\begin{aligned}|\Delta_\omega(x)| &\leq C_+2^{s+1}(1+\|\omega\|_2^2)^{-s/2-1/2} \cdot \|\Sigma(x) - I\|_{\text{op}}\|\omega\|_2 \\ &\leq C_+2^{s+1}(1+\|\omega\|_2^2)^{-s/2} \cdot \|\Sigma(x) - I\|_{\text{op}}\end{aligned}$$

By Assumption A.7(3),

$$\|\Delta_\omega\|_{L^1} \leq \psi C_+ 2^{s+1}(1+\|\omega\|_2^2)^{-s/2}$$

By Assumption A.7(2) and Lemma A.6, for any $\alpha \in \mathbb{N}^J$ with $0 < |\alpha| \leq M$

$$\|D_x^\alpha \Delta_\omega\|_{L^1} = \|D_x^\alpha \hat{q}(\Sigma(x)\omega)\|_{L^1} \leq \psi C_+ C_{J,s}(1+\|\omega\|_2^2)^{-s/2}.$$

Step 2: Bound the Fourier transform $|\hat{\Delta}_\omega|$ with L^1 norm of derivatives of Δ_ω.

By Proposition A.8,

$$\begin{aligned}|\hat{\Delta}_\omega(\upsilon)| &\leq \psi\tilde{C}(1+\|\omega\|_2^2)^{-s/2}(1+\|\upsilon\|_2)^{-M} \\ &\leq \psi\tilde{C}(1+\|\omega\|_2^2)^{-s/2}(1+\|\upsilon\|_2^2)^{-M/2},\end{aligned}$$

where

$$\tilde{C} = c_{J,M}C_+C_{J,s}.$$

Step 3: Bound the key condition (A.17).

By Assumption A.7(2),

$$\begin{aligned}&\left|\int_{\mathbb{R}^J} |\hat{\Delta}_\omega(\omega-\upsilon)|\frac{\hat{q}(\omega)}{\hat{q}^2(\upsilon)}d\omega\right| \\ &\leq \psi\frac{\tilde{C}C_+}{C_-^2}\int_{\mathbb{R}^J}\frac{(1+\|\omega-\upsilon\|_2^2)^{-M/2}(1+\|\omega\|_2^2)^{-s}}{(1+\|\upsilon\|_2^2)^{-s}}d\omega\end{aligned}$$

$$= \psi \frac{\tilde{C} C_+}{C_-^2} \int_{\mathbb{R}^J} \frac{(1+\|\upsilon\|_2^2)^s}{(1+\|\omega-\upsilon\|_2^2)^{M/2}(1+\|\omega\|_2^2)^s} d\omega.$$

Note that

$$1+\|\upsilon\|_2^2 \leq 1+2(\|\omega\|_2^2+\|\omega-\upsilon\|_2^2) \leq 2(1+\|\omega\|_2^2)(1+\|\omega-\upsilon\|_2^2).$$

Thus,

$$\begin{aligned}
&\left|\int_{\mathbb{R}^J} |\hat{\Delta}_\omega(\omega-\upsilon)| \frac{\hat{q}(\omega)}{\hat{q}^2(\upsilon)} d\omega\right| \\
&\leq \psi \frac{2^s \tilde{C} C_+}{C_-^2} \int_{\mathbb{R}^J} (1+\|\omega-\upsilon\|_2^2)^{-(\frac{M}{2}-s)} d\omega \\
&= \psi \frac{2^s \tilde{C} C_+}{C_-^2} \int_{\mathbb{R}^J} (1+\|\upsilon\|_2^2)^{-(\frac{M}{2}-s)} d\omega \\
&= \psi \frac{2^s \tilde{C} C_+}{C_-^2} \int_{\mathbb{R}^J} (1+\|\upsilon\|_2^2)^{-(J+1)/2} d\omega \\
&= \psi \frac{2^s \tilde{C} C_+}{C_-^2} \frac{\pi^{(J+1)/2}}{\Gamma((J+1)/2)}.
\end{aligned}$$

This can be made less than 1 because we can choose

$$\psi < \tilde{\psi}^*(J, s, C_\pm) = \frac{C_-^2}{2^s c_{J,M} C_+ C_{J,s} C_+} \frac{\Gamma((J+1)/2)}{\pi^{(J+1)/2}}. \tag{A.13}$$

The bound is uniform in ω. Thus, (A.17) is proved. The proof is then completed. □

Lemma A.4. *For $g : \mathbb{R}^J \mapsto \mathbb{R}^J$, let $\mathcal{T}$ be the operator defined by*

$$(\mathcal{T}_0 g)(x) \equiv \int_{\mathbb{R}^J} q(x-\delta) g(\delta) d\delta.$$

Under Assumption A.7(2) and Assumption A.7(3), $\mathbb{W}^{2s,\infty}$ has a unique preimage in $\mathbb{W}^{s,\infty}$ under $\mathcal{T}_0$.

Proof. By Lemma A.8,

$$(\mathcal{T}_0 g)(x) = (2\pi)^{-J} \int_{\mathbb{R}^J} e^{i\langle x,\omega\rangle} \hat{q}(\omega) \hat{g}(\omega) d\omega. \tag{A.14}$$

Note that $\mathcal{T}_0$ is a multiplication operator, i.e.,

$$\widehat{\mathcal{T}_0 g}(\omega) = \hat{q}(\omega)\hat{g}(\omega).$$

For any $k \in \mathbb{W}^{2s,\infty}$, let u be the function with

$$\hat{u}(\omega) = \frac{\hat{k}(\omega)}{\hat{q}(\omega)}. \tag{A.15}$$

Assumption A.7(2) implies

$$|\hat{u}(\omega)| = \underbrace{\frac{|\hat{k}(\omega)|}{|\hat{q}^2(\omega)|}}_{\in L^\infty(\mathbb{R}^J)} \cdot \underbrace{|\hat{q}(\omega)|}_{\in \mathbb{W}^{s,\infty}}.$$

Thus, $u \in \mathbb{W}^{s,\infty}$. This also implies that u is a tempered distribution (which can be identified with a function). By Theorem 7.1.10 of Hörmander (2003),

$$\widehat{\mathcal{T}_0 u} = \hat{k} \Longleftrightarrow \mathcal{T}_0 u = k.$$

Thus,

$$\mathcal{T}_0^{-1} k \in \mathbb{W}^{s,\infty}, \quad \forall k \in \mathbb{W}^{2s,\infty}.$$

□

Lemma A.5. *Let $\mathcal{T}$ be the operator*

$$(\mathcal{T} g)(x) = \int_{\mathbb{R}^J} q\left(\Sigma(x)^{-1}(x-t)\right) g(t)\, dt = \mathbb{E}_{\delta = X + \Sigma(X)\xi, \xi \sim q(\cdot)}[g(\delta) \mid X = x]. \tag{A.16}$$

For any $\omega \in \mathbb{R}^J$, let

$$\Delta_\omega(x) = \hat{q}(\Sigma(x)\omega) - \hat{q}(\omega).$$

Assume that

$$\max_{\upsilon \in \mathbb{R}^J} \left| \int_{\mathbb{R}^J} |\hat{\Delta}_\omega(\omega - \upsilon)| \frac{\hat{q}(\omega)}{\hat{q}^2(\upsilon)} d\omega \right| \leq 1 - \eta \tag{A.17}$$

for some constant $\eta > 0$. Under Assumption A.7(1)–(3), $\mathbb{W}^{2s,\infty}$ has a unique preimage in $\mathbb{W}^{s,\infty}$ under $\mathcal{T}$.

Proof. Define $\mathcal{E} = \mathcal{T} - \mathcal{T}_0$. By Lemma A.8 and (A.14), for any $g \in \mathbb{W}^{s,\infty}$,

$$(\mathcal{E} g)(x) = (2\pi)^{-J} \int_{\mathbb{R}^J} e^{i\langle x, \omega\rangle} \Delta_\omega(x) \hat{g}(\omega) d\omega.$$

Then,

$$\begin{aligned} \widehat{(\mathcal{E} g)}(\upsilon) &= \int_{\mathbb{R}^J} e^{-i\langle x, \upsilon\rangle} \left((2\pi)^{-J} \int_{\mathbb{R}^J} e^{i\langle x, \omega\rangle} \Delta_\omega(x) \hat{g}(\omega) d\omega \right) dx \\ &= \int_{\mathbb{R}^J} \int_{\mathbb{R}^J} (2\pi)^{-J} e^{i\langle x, \omega - \upsilon\rangle} \Delta_\omega(x) \hat{g}(\omega) dx d\omega \\ &= \int_{\mathbb{R}^J} \hat{\Delta}_\omega(\omega - \upsilon) \hat{g}(\omega) d\omega. \end{aligned} \tag{A.18}$$

By Assumption A.7(2) and the standard definition of $\mathbb{W}^{s,\infty}$, for any $k \in \mathbb{W}^{s,\infty}(\mathbb{R}^J)$,

$$\|k\|_{\mathbb{W}^{s,\infty}} < \infty \Longleftrightarrow \left\| \frac{\hat{k}}{\hat{q}} \right\|_{L^\infty} \equiv \|k\|_{\hat{q}} < \infty.$$

Then we have

$$\mathbb{W}^{s,\infty}(\mathbb{R}^J) = \left\{ k : \|k\|_{\hat{q}} < \infty \right\}.$$

By (A.15), we have that $\widehat{\mathcal{T}_0^{-1}\rho} = \hat{\rho}/\hat{q}$ for any $\rho \in \mathbb{W}^{2s,\infty}$. We first verify that $\mathcal{E}g \in \mathbb{W}^{2s,\infty}$. By (A.18),

$$\begin{aligned}
\sup_{\upsilon\in\mathbb{R}^J} \frac{\widehat{(\mathcal{E}g)}(\upsilon)}{|\hat{q}(\upsilon)|^2} &= \sup_{\upsilon\in\mathbb{R}^J} \int_{\mathbb{R}^J} \frac{|\hat{\Delta}_\omega(\omega-\upsilon)\hat{g}(\omega)|}{|\hat{q}^2(\upsilon)|} d\omega \\
&= \sup_{\upsilon\in\mathbb{R}^J} \int_{\mathbb{R}^J} |\hat{\Delta}_\omega(\omega-\upsilon)| \frac{|\hat{q}(\omega)|}{|\hat{q}(\upsilon)|^2} \frac{|\hat{g}(\omega)|}{|\hat{q}(\omega)|} d\omega \\
&= \sup_{\upsilon\in\mathbb{R}^J} \int_{\mathbb{R}^J} |\hat{\Delta}_\omega(\omega-\upsilon)| \frac{|\hat{q}(\omega)|}{|\hat{q}(\upsilon)|^2} d\omega \cdot \left\| \frac{\hat{g}}{\hat{q}} \right\|_{L^\infty} \\
&\leq (1-\eta)\|g\|_{\hat{q}} \qquad \text{(A.19)} \\
&< \infty,
\end{aligned}$$

where (A.19) uses the condition (A.17) and the last line follows from $g \in \mathbb{W}^{s,\infty}$. Now, we apply (A.15):

$$\begin{aligned}
\|\mathcal{T}_0^{-1}\mathcal{E}g\|_{\hat{q}} &= \left\| \frac{\widehat{(\mathcal{T}_0^{-1}\mathcal{E}g)}}{\hat{q}} \right\|_{L^\infty} \\
&= \sup_{\upsilon\in\mathbb{R}^J} \int_{\mathbb{R}^J} \frac{|\hat{\Delta}_\omega(\omega-\upsilon)\hat{g}(\omega)|}{|\hat{q}^2(\upsilon)|} d\omega \\
&\leq (1-\eta)\|g\|_{\hat{q}},
\end{aligned}$$

where the last line follows from (A.19). This implies that

$$\|\mathcal{T}_0^{-1}\mathcal{E}\|_{\|\cdot\|_{\hat{q}} \leftrightarrow \|\cdot\|_{\hat{q}}} = \|\mathcal{T}_0^{-1}\mathcal{E}\|_{\|\cdot\|_{\mathbb{W}^{s,\infty}} \leftrightarrow \|\cdot\|_{\mathbb{W}^{s,\infty}}} \leq 1-\eta.$$

As a result, the following Neumann series converges on $\mathbb{W}^{s,\infty}$:

$$\sum_{\ell\geq 0} (-\mathcal{T}_0^{-1}\mathcal{E})^\ell.$$

By Lemma A.4, for any $k \in \mathbb{W}^{2s,\infty}$, $\mathcal{T}_0^{-1}k \in \mathbb{W}^{s,\infty}$. Let

$$g = \sum_{\ell\geq 0} (-\mathcal{T}_0^{-1}\mathcal{E})^\ell \mathcal{T}_0^{-1}k.$$

Then $\|g\|_{\mathbb{W}^{s,\infty}} \leq \eta^{-1} \|\mathcal{T}_0^{-1} k\|_{\mathbb{W}^{s,\infty}} < \infty$ and

$$(\mathrm{Id} + \mathcal{T}_0^{-1}\mathcal{E})g = \mathcal{T}_0^{-1}k.$$

We thus have that

$$\mathcal{T}g = (\mathcal{T}_0 + \mathcal{E})g = \mathcal{T}_0(\mathrm{Id} + \mathcal{T}_0^{-1}\mathcal{E})g = \mathcal{T}_0\mathcal{T}_0^{-1}k = k.$$

□

Lemma A.6. *Under Assumption A.7(2)–(3), for any* $\alpha \in \mathbb{N}^J$ *with* $|\alpha| \leq M$ *and* $\omega \in \mathbb{R}^J$,

$$\sup_{x \in \mathbb{R}^J} |D_x^\alpha \hat{q}(\Sigma(x)\omega)| + \int_{\mathbb{R}^J} |\partial_x^\alpha \hat{q}(\Sigma(x)\omega)| dx \leq \psi C_+ C_{J,s} (1 + \|\omega\|_2^2)^{-s/2},$$

where

$$C_{J,s} = 2^{s+1+M} \sum_{\kappa \in \mathbb{N}^J, 1 \leq |\kappa| \leq |\alpha|} \sum_{\substack{(k_{\gamma,j}) \in \mathbb{N} \\ \sum_\gamma k_{\gamma,j} = \kappa_j, \forall j \\ \sum_{\gamma,j} k_{\gamma,j}\gamma = \alpha}} \frac{\alpha!}{\prod_\gamma k_{\gamma,j}! \, (\gamma!)^{k_{\gamma,j}}}.$$

Proof. Fix $\omega \in \mathbb{R}^J$. Let $g(x) = \Sigma(x)\omega$. For any γ with $1 \leq |\gamma| \leq |\alpha|$,

$$D_x^\gamma g(x) = (D_x^\gamma \Sigma(x))\, \omega.$$

Thus,

$$\|D_x^\gamma g(x)\|_{\mathrm{op}} \leq \|D_x^\gamma \Sigma(x)\|_{\mathrm{op}} \|\omega\|$$

Furthermore, by Assumption A.7(3),

$$|D_v^\kappa \hat{q}(\Sigma(x)\omega)| \leq C_+ (1 + \|\Sigma(x)\omega\|_2^2)^{-s/2 - |\kappa|/2}.$$

By Assumption A.7(2),

$$\|\Sigma(x)\|_{\mathrm{op}} \geq 1 - \psi \geq \frac{1}{2}.$$

Then

$$\begin{aligned}
|D_v^\kappa \hat{q}(\Sigma(x)\omega)| &\leq C_+ (1 + \|\omega\|_2^2/4)^{-s/2 - |\kappa|/2} \\
&\leq C_+ 2^{s+|\kappa|} (1 + \|\omega\|_2^2)^{-s/2 - |\kappa|/2} \\
&\leq C_+ 2^{s+|\alpha|} (1 + \|\omega\|_2^2)^{-s/2 - |\kappa|/2} \\
&\triangleq C_1 (1 + \|\omega\|_2^2)^{-s/2 - |\kappa|/2}
\end{aligned}$$

By the multivariate Faà di Bruno formula (Proposition A.7),

$|D_x^\alpha \hat{q}(\Sigma(x)\omega)|$

$$
\leq \sum_{\kappa\in\mathbb{N}^J,1\leq|\kappa|\leq|\alpha|} |D_v^\kappa \hat{q}(\Sigma(x)\omega)| \sum_{\substack{(k_{\gamma,j})\in\mathbb{N} \\ \sum_\gamma k_{\gamma,j}=\kappa_j,\ \forall j \\ \sum_{\gamma,j} k_{\gamma,j}\,\gamma=\alpha}} \frac{\alpha!}{\prod_{\gamma,j} k_{\gamma,j}!\,(\gamma!)^{k_{\gamma,j}}} \prod_{\gamma,j} |D_x^\gamma g_j(x)|^{k_{\gamma,j}}
$$

$$
\leq C_+ 2^{s+|\alpha|} \sum_{\kappa\in\mathbb{N}^J,1\leq|\kappa|\leq|\alpha|} (1+\|\omega\|_2^2)^{-s-|\kappa|/2}
$$

$$
\sum_{\substack{(k_{\gamma,j})\in\mathbb{N} \\ \sum_\gamma k_{\gamma,j}=\kappa_j,\forall j \\ \sum_{\gamma,j} k_{\gamma,j}\,\gamma=\alpha}} \frac{\alpha!}{\prod_{\gamma,j} k_{\gamma,j}!\,(\gamma!)^{k_{\gamma,j}}} \prod_{\gamma,j} (\|D_x^\gamma \Sigma(x)\|_{\text{op}} \|\omega\|_2)^{k_{\gamma,j}}
$$

$$
\leq C_+ 2^{s+|\alpha|} (1+\|\omega\|_2^2)^{-s} \sum_{\kappa\in\mathbb{N}^J,1\leq|\kappa|\leq|\alpha|} (1+\|\omega\|_2^2)^{-|\kappa|/2} \|\omega\|_2^{|\kappa|}
$$

$$
\cdot \sum_{\substack{(k_{\gamma,j})\in\mathbb{N} \\ \sum_\gamma k_{\gamma,j}=\kappa_j,\forall j \\ \sum_{\gamma,j} k_{\gamma,j}\,\gamma=\alpha}} \frac{\alpha!}{\prod_{\gamma} k_{\gamma,j}!\,(\gamma!)^{k_{\gamma,j}}} \prod_{\gamma,j} \|D_x^{\gamma,j} \Sigma(x)\|_{\text{op}}^{k_{\gamma,j}}
$$

$$
\leq C_+ 2^{s+|\alpha|} (1+\|\omega\|_2^2)^{-s} \sum_{\kappa\in\mathbb{N}^J,1\leq|\kappa|\leq|\alpha|} \sum_{\substack{(k_{\gamma,j})\in\mathbb{N} \\ \sum_\gamma k_{\gamma,j}=\kappa_j,\forall j \\ \sum_{\gamma,j} k_{\gamma,j}\,\gamma=\alpha}} \frac{\alpha!}{\prod_{\gamma,j} k_{\gamma,j}!\,(\gamma!)^{k_{\gamma,j}}} \prod_{\gamma,j} \|D_x^\gamma \Sigma(x)\|_{\text{op}}^{k_{\gamma,j}}.
$$

Let

$$
C_1 = \sum_{\kappa\in\mathbb{N}^J,1\leq|\kappa|\leq|\alpha|} \sum_{\substack{(k_{\gamma,j})\in\mathbb{N} \\ \sum_\gamma k_{\gamma,j}=\kappa_j,\forall j \\ \sum_{\gamma,j} k_{\gamma,j}\,\gamma=\alpha}} \frac{\alpha!}{\prod_{\gamma} k_{\gamma,j}!\,(\gamma!)^{k_{\gamma,j}}}.
$$

By Assumption A.7(3),

$$
\prod_{\gamma,j} \|D_x^\gamma \Sigma(x)\|_{\text{op}}^{k_{\gamma,j}} \leq \psi^{|\kappa|} \leq \psi.
$$

Thus,

$$
|D_x^\alpha \hat{q}(\Sigma(x)\omega)| \leq \psi C_+ 2^{s+|\alpha|} C_1 (1+\|\omega\|_2^2)^{-s/2}.
$$

Similarly, to bound the L^1 norm of $D_x^\alpha \hat{q}(\Sigma(x)\omega)$, we just need to show that

$$
\left\| \prod_{\gamma,j} \|D_x^\gamma \Sigma(x)\|_{\text{op}}^{k_{\gamma,j}} \right\|_{L^1} \leq \psi.
$$

This is because at least one $k_\gamma \geq 1$ and $\|ab\|_{L^1} \leq \|a\|_{L^1}\|b\|_{L^\infty}$. The proof is then completed by setting $C_{J,s} = 2^{s+|\alpha|+1} C_+ C_1$. □

A.7.4. *Auxiliary lemmas.*

Lemma A.7. *If $\hat{q}(\upsilon) = (1 + \|\upsilon\|_2^2)^{-s/2}$ for some $s > J$, then Assumption A.7(2) holds.*

Remark A.3. Since $\hat{q}(0) = 1$, $\hat{q}$ is the characteristic function of a density. In fact, for any $s > J$, q has the following expression

$$q(x) = \Theta_{J,s} \|x\|_2^{(s-J)/2} K_{(s-J)/2}(\|x\|_2)$$

where K_v is the modified Bessel function of the second kind and $\Theta_{J,s}$ is the normalizing constant. For large x,

$$q(x) \sim \|x\|_2^{(s-J-1)/2} \exp\{-\|x\|_2\}.$$

Thus, q behaves like a radial Gamma distribution.

Proof. We will prove that

$$D_\upsilon^\alpha \hat{q}(\upsilon) = \sum_{\ell=1}^{|\alpha|} (1 + \|\upsilon\|_2^2)^{-s/2-\ell} F_{\alpha,\ell}(\upsilon),$$

where $F_{\alpha,\ell}$ is a homogeneous polynomial of υ of order ℓ. We prove this claim by induction on $|\alpha|$. When $\ell = 1$,

$$D_\upsilon^{e_j} \hat{q}(\upsilon) = (1 + \|\upsilon\|_2^2)^{-s/2-1} \cdot 2\upsilon_j$$

when e_j is the j-th canonical basis. Suppose the claim holds for $|\alpha| - 1$. For any given α, assume WLOG that $\alpha_1 \geq 1$. Let $\tilde{\alpha} = \alpha - e_1$. Then

$$\begin{aligned}
D_\upsilon^\alpha \hat{q}(\upsilon) = D_\upsilon^{e_1} D_\upsilon^{\tilde{\alpha}} \hat{q}(\upsilon) &= D_\upsilon^{e_1} \sum_{\ell=1}^{|\tilde{\alpha}|} (1 + \|\upsilon\|_2^2)^{-s/2-\ell} F_{\tilde{\alpha},\ell}(\upsilon) \\
&= \sum_{\ell=1}^{|\tilde{\alpha}|} (1 + \|\upsilon\|_2^2)^{-s/2-\ell-1} \cdot -(s + 2\ell)\upsilon_1 F_{\tilde{\alpha},\ell}(\upsilon) \\
&\quad + \sum_{\ell=1}^{|\tilde{\alpha}|} (1 + \|\upsilon\|_2^2)^{-s/2-\ell} \cdot D_\upsilon^{e_1} F_{\tilde{\alpha},\ell}(\upsilon) \\
&= \sum_{\ell=1}^{|\tilde{\alpha}|} (1 + \|\upsilon\|_2^2)^{-s/2-\ell} \cdot \left(D_\upsilon^{e_1} F_{\tilde{\alpha},\ell}(\upsilon) - (s + 2\ell - 1)\upsilon_1 F_{\tilde{\alpha},\ell-1}(\upsilon) \right).
\end{aligned}$$

Clearly, $D_\upsilon^{e_1} F_{\tilde{\alpha},\ell}(\upsilon) - (s + 2\ell - 1)\upsilon_1 F_{\tilde{\alpha},\ell-1}(\upsilon)$ is a homogeneous polynomial of order ℓ. Thus, the claim holds for $|\alpha|$. By induction the proof is completed. □

Proposition A.7 (Multivariate Faà di Bruno formula; see e.g., Theorem 6.8 of Schumann (2019))**.** *Let $g(x) = (g_1(x), \ldots, g_n(x)) : \mathbb{R}^m \mapsto \mathbb{R}^n$ and $f(y) : \mathbb{R}^n \mapsto \mathbb{R}$ be two* *functions that have M derivatives over $\mathbb{R}^n$. Then*

$$D_x^\alpha (f \circ g)(x) = \sum_{\kappa \in \mathbb{N}^n, 1 \le |\kappa| \le |\alpha|} D_y^\kappa f(g(x)) B_{\alpha,\kappa}((D_x^\gamma g_j(x))_{\gamma \in \mathbb{N}^n \setminus \{\mathbf{0}\}, 1 \le j \le n}),$$

where $B_{\alpha,\kappa}$ is the multivariate Bell polynomial defined as

$$B_{\alpha,\kappa}\big((z_{\gamma,j})_{\gamma \in \mathbb{N}^n \setminus \{\mathbf{0}\},\, 1 \le j \le n}\big) = \sum_{\substack{(k_{\gamma,j}) \in \mathbb{N} \\ \sum_\gamma k_{\gamma,j} = \kappa_j \ \forall j \\ \sum_{\gamma,j} k_{\gamma,j}\, \gamma = \alpha}} \frac{\alpha!}{\prod_{\gamma,j} k_{\gamma,j}!\, (\gamma!)^{k_{\gamma,j}}} \prod_{\gamma,j} z_{\gamma,j}^{k_{\gamma,j}}.$$

Proposition A.8 (Corollary 3.3.10 of Grafakos et al. (2008))**.** *Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ with $D^\alpha f \in L^1$ for any $\alpha \in \mathbb{N}^n$ with $|\alpha| \le M$. Then, for any $\upsilon \in \mathbb{R}^J$,*

$$|\hat{f}(\upsilon)| \le c_{n,M} \max\left\{ \|f\|_{L^1}, \max_{|\alpha|=M} \|D^\alpha f\|_{L^1} \right\} (1 + \|\upsilon\|)^{-M},$$

where $c_{n,M}$ is a constant that only depends on n and M.

Lemma A.8. *Let q be a density that satisfies Assumption A.7(2). For some $\Sigma(x)$ taking values in $\mathbb{S}_+^J$, let $\mathcal{T}$ be defined as in (A.16) and let $q_x(\zeta) = \det(\Sigma(x))^{-1} q\,(\Sigma(x)^{-1}\zeta)$. Then, $\mathcal{T}$ is a pseudo-differential operator with symbol $\hat{q}_x(\omega)$ in Kohn–Nirenberg quantization, i.e.,*

$$(\mathcal{T} g)(x) = \frac{1}{(2\pi)^J} \int_{\mathbb{R}^J} e^{i\langle x, \omega\rangle} \hat{q}_x(\omega) \hat{g}(\omega) d\omega.$$

Proof. By the assumptions on $q(\cdot)$,

$$\begin{aligned}
(\mathcal{T} g)(x) &= \int_{\mathbb{R}^J} q_x(x - \delta) g(\delta) d\delta \\
&= \frac{1}{(2\pi)^J} \int_{\mathbb{R}^J} \int_{\mathbb{R}^J} e^{i\langle x-\delta, \omega\rangle} \hat{q}_x(\omega) g(\delta) d\omega d\delta \qquad \text{(Inverse Fourier transform)} \\
&= \frac{1}{(2\pi)^J} \int_{\mathbb{R}^J} \int_{\mathbb{R}^J} e^{i\langle x, \omega\rangle} \hat{q}_x(\omega) e^{-i\langle \delta, \omega\rangle} g(\delta) d\delta d\omega \\
&= \frac{1}{(2\pi)^J} \int_{\mathbb{R}^J} \int_{\mathbb{R}^J} e^{i\langle x, \omega\rangle} \hat{q}_x(\omega) \hat{g}(\omega) d\omega.
\end{aligned}$$

□

Lemma A.9. *Assume $M \ge 2s \ge 1$. If $k \in \mathbb{K}^{2s}$, then and $k \circ b^{-1} \in \mathbb{K}^{2s}$.*

Proof. By (A.20), $b^{-1} \in \mathbb{K}^M$. Since $M \ge 2s$, Lemma A.12 implies $k \circ b^{-1} \in \mathbb{K}^{2s}$. □

Lemma A.10. *Under Assumption A.7,*

$$\|\tilde{\Sigma}(\cdot) - I\|_{\mathbb{W}^{M,\infty}} + \|\tilde{\Sigma}(\cdot) - I\|_{\mathbb{W}^{M,1}}$$

$$\leq B_M\Big(\|\Sigma(\cdot)-\Sigma_0\|_{\mathbb{W}^{M,\infty}}+\|\Sigma(\cdot)-\Sigma_0\|_{\mathbb{W}^{M,1}},\|Db\|_{\mathbb{W}^{M-1,\infty}},\|Db^{-1}\|_{L^\infty},\lambda_{\min}^{-1}(\Sigma_0)\Big)$$

for some function B_M that only depends on M. In particular,

$$\lim_{y\to 0} B_M\left(y,\|Db\|_{\mathbb{W}^{M-1,\infty}},\|Db^{-1}\|_{L^\infty},\lambda_{\min}^{-1}(\Sigma_0)\right)=0.$$

Proof. First, we note that

$$\begin{aligned}
&\|\tilde{\Sigma}(\cdot)-I\|_{\mathbb{W}^{M,\infty}}+\|\tilde{\Sigma}(\cdot)-I\|_{\mathbb{W}^{M,1}}\\
&\leq \left(\|\Sigma\circ b^{-1}-\Sigma_0\|_{\mathbb{W}^{M,\infty}}+\|\Sigma\circ b^{-1}-\Sigma_0\|_{\mathbb{W}^{M,1}}\right)\lambda_{\min}^{-1}(\Sigma_0)\\
&=\left(\|(\Sigma(\cdot)-\Sigma_0)\circ b^{-1}\|_{\mathbb{W}^{M,\infty}}+\|(\Sigma(\cdot)-\Sigma_0)\circ b^{-1}\|_{\mathbb{W}^{M,1}}\right)\lambda_{\min}^{-1}(\Sigma_0)
\end{aligned}$$

By Assumption B3, $\Sigma(\cdot)-\Sigma_0\in\mathbb{W}^{M,\infty}$. By Lemma A.11 and Assumption A.7(4),

$$b^{-1}\in\mathbb{K}^M \implies \|Db^{-1}\|_{\mathbb{W}^{M-1,\infty}}<\infty. \tag{A.20}$$

Furthermore,

$$b\in\mathbb{K}^M \implies Db\in L^\infty. \tag{A.21}$$

By Lemma A.13, (A.20), and (A.21),

$$\left\|(\Sigma(\cdot)-\Sigma_0)\circ b^{-1}\right\|_{\mathbb{W}^{M,\infty}}\leq B_{M,\infty}(\|\Sigma(\cdot)-\Sigma_0\|_{\mathbb{W}^{M,\infty}},\|Db^{-1}\|_{\mathbb{W}^{M-1,\infty}},\|Db\|_{L^\infty}),$$

and

$$\left\|(\Sigma(\cdot)-\Sigma_0)\circ b^{-1}\right\|_{\mathbb{W}^{M,1}}\leq B_{M,1}(\|\Sigma(\cdot)-\Sigma_0\|_{\mathbb{W}^{M,1}},\|Db^{-1}\|_{\mathbb{W}^{M-1,\infty}},\|Db\|_{L^\infty}).$$

The proof is completed by defining

$$\begin{aligned}
&B_M\Big(\|\Sigma(\cdot)-\Sigma_0\|_{\mathbb{W}^{M,\infty}}+\|\Sigma(\cdot)-\Sigma_0\|_{\mathbb{W}^{M,1}},\|Db\|_{\mathbb{W}^{M-1,\infty}},\|Db^{-1}\|_{L^\infty},\|\Sigma_0^{-1}\|_{\mathrm{op}}\Big)\\
&=\Big\{B_{M,\infty}(\|\Sigma(\cdot)-\Sigma_0\|_{\mathbb{W}^{M,\infty}},\|Db^{-1}\|_{\mathbb{W}^{M-1,\infty}},\|Db\|_{L^\infty})\\
&\qquad+B_{M,1}(\|\Sigma(\cdot)-\Sigma_0\|_{\mathbb{W}^{M,1}},\|Db^{-1}\|_{\mathbb{W}^{M-1,\infty}},\|Db\|_{L^\infty})\Big\}\lambda_{\min}^{-1}(\Sigma_0).
\end{aligned}$$

□

Lemma A.11. *Let $s\geq 1$ and let $f(x)=Ax+u(x)\in\mathbb{K}^s$. Assume*

$$\left\|(Df)^{-1}\right\|_{L^\infty}<\infty. \tag{A.22}$$

Then $f^{-1}\in\mathbb{K}^s$

Proof. We split the proof into a few steps.

Step 1: the linear part A is invertible

To contradiction, if A were not full rank, then there exists $0 \neq e \in \mathrm{Ran}(A)^\perp$ (equivalently $A^T e = 0$). For all $x \in \mathbb{R}^J$,

$$e \cdot f(x) = e \cdot (Ax + u(x)) = e \cdot u(x).$$

Since $u \in \mathbb{W}^{s,\infty} \subset L^\infty$, the right-hand side is bounded; hence $e \cdot f(\mathbb{R}^J)$ is bounded. If f is surjective, $f(\mathbb{R}^J) = \mathbb{R}^J$, but $e \cdot y$ is unbounded on $\mathbb{R}^J$, a contradiction. Thus A is full rank.

Step 2: isolate the linear part of the inverse
Let $g = f^{-1}$. Using $y = f(g(y)) = Ag(y) + u(g(y))$ we obtain

$$g(y) = A^{-1}y - A^{-1}u(g(y)) =: A^{-1}y + v(y), \qquad v(y) := -A^{-1}u(g(y)).$$

Since $u \in L^\infty$, we have $v \in L^\infty$ and

$$\|v\|_{L^\infty} \leq \|A^{-1}\| \, \|u\|_{L^\infty}.$$

Hence $g \in \mathrm{lin} \oplus L^\infty$. It remains to show $v \in \mathbb{W}^{s,\infty}$.

Step 3: prove the result for integer s
Differentiating $f(g(y)) = y$ gives

$$Df(g(y))\, Dg(y) = I, \tag{A.23}$$

Then

$$Dg(y) = \big(Df(g(y))\big)^{-1}.$$

The assumption $\|(Df)^{-1}\|_{L^\infty} < \infty$ implies $Dg \in L^\infty$. Since $g(y) = A^{-1}y + v(y)$, we have $Dv = Dg - A^{-1} \in L^\infty$. Thus, $v \in \mathbb{W}^{1,\infty}$.

We now prove $v \in \mathbb{W}^{m,\infty}$ for all $1 \leq m \leq s$ by induction. Suppose for some $2 \leq m \leq s$ we prove that $v \in \mathbb{W}^{m-1,\infty}$. Then

$$D^j v \in L^\infty, \quad j = 1, \dots, m-1. \tag{A.24}$$

Differentiating (A.23) m times we obtain, by the Leibniz rule and the multivariate Faà di Bruno formula,

$$\{(Df) \circ g\} \cdot D^m g = R_m((D^1 f) \circ g, \dots, (D^m f) \circ g, D^1 g, \dots, D^{m-1} g),$$

where $D^k g, (D^k f) \circ g \in (\mathbb{R}^J)^{\otimes k}$ is the k-th differentials and R_m is a tensor in $(\mathbb{R}^J)^{\otimes m}$ for which each coordinate is a polynomial of the entries of $(D^1 f) \circ g, \dots, (D^m f) \circ g, D^1 g, \dots, D^{m-1} g$. Since $f \in \mathbb{K}^s$, $D^k \in L^\infty$. Thus, by the induction hypothesis

(A.24), $R_m \in L^\infty$. By (A.22),

$$D^m g = (Df(g))^{-1} R_m \in L^\infty.$$

Since $Dv = Dg - A^{-1}$, $D^m v = D^m g \in L^\infty$. Thus, $v \in \mathbb{W}^{m,\infty}$. The induction argument then implies $v \in \mathbb{W}^{s,\infty}$.

Step 4: prove the result for non-integer s

Let $s = m + \sigma$ with $m = \lfloor s \rfloor \in \mathbb{N}$ and $\sigma \in (0, 1)$. For any $1 \le k \le m$, $f \in \mathbb{K}^s$ implies

$$D^k f \in \mathbb{W}^{\sigma,\infty}.$$

In Step 3, we have proved that $g \in \mathbb{W}^{m,\infty} \in \mathbb{W}^{1,\infty}$ and thus g is Lipschitz. By Lemma A.14,

$$(D^k f) \circ g \in \mathbb{W}^{\sigma,\infty}, \quad k = 1, \ldots, m.$$

By (A.22) and Lemma A.15,

$$((Df) \circ g)^{-1} \in \mathbb{W}^{\sigma,\infty}.$$

Recall that each coordinate of R_m is a polynomial of the entries of $(D^1 f) \circ g, \ldots, (D^m f) \circ g, D^1 g, \ldots, D^{m-1} g$. Since $g \in \mathbb{W}^{m,\infty}$, $D^1 g, \ldots, D^{m-1} g$ are all Lipschitz and hence in $\mathbb{W}^{\sigma,\infty}$. Since each monomial is a product, Lemma A.16 implies

$$R_m \in \mathbb{W}^{\sigma,\infty}.$$

By Lemma A.16 again, we conclude that

$$D^m v = D^m g - A^{-1} I(m = 1) \in \mathbb{W}^{\sigma,\infty}.$$

Therefore, $g \in \mathbb{K}^s$. □

Lemma A.12. *If $f, g \in \mathbb{K}^s$ for some $s \ge 1$, then $f \circ g \in \mathbb{K}^s$.*

Proof. We split the proof into a few steps.

Step 1: reduction to the nonlinear part

Let

$$f(x) = Ax + u(x), \qquad g(x) = Bx + v(x),$$

with $u, v \in \mathbb{W}^{s,\infty}$. Then

$$(f \circ g)(x) = A(Bx + v(x)) + u(Bx + v(x)) = (AB)x + Av(x) + u \circ g(x).$$

Since $Av \in \mathbb{W}^{s,\infty}$, it suffices to show $u \circ g \in \mathbb{W}^{s,\infty}$.

Step 2: proof for integer s

Since $g \in \mathbb{K}^s$, $D^k g \in \mathbb{W}^{s-k,\infty}$. By the multivariate Faà di Bruno formula, for each integer $k \leq s$,

$$D^k(u \circ g) = S_k((D^1 u) \circ g, \ldots, (D^k u) \circ g, D^1 g, \ldots, D^k g), \tag{A.25}$$

where S_k is a tensor in $(\mathbb{R}^J)^{\otimes k}$ for which each coordinate is a polynomial of the entries of $(D^1 f) \circ g, \ldots, (D^k f) \circ g, D^1 g, \ldots, D^k g$. In particular, each monomial involves at least one coordinate of $D^j f \circ g$ for some j.

Since $k \leq s$,

$$(D^j u) \circ g \in L^\infty, \; D^j g \in L^\infty, \quad j = 1, \ldots, k.$$

Thus,

$$D^k(u \circ g) \in L^\infty.$$

Since this holds for all $k \leq s$, we conclude that $u \circ g \in \mathbb{W}^{s,\infty}$.

Step 3: proof for non-integer s

Let $s = m + \sigma$ with $m = \lfloor s \rfloor \in \mathbb{N}$ and $\sigma \in (0,1)$. In Step 2, we have already proved that

$$D^m(u \circ g) = S_m((D^1 u) \circ g, \ldots, (D^m u) \circ g, D^1 g, \ldots, D^m g).$$

Since $u, g \in \mathbb{W}^{m+\sigma,\infty}$,

$$(D^1 u) \circ g, \ldots, (D^m u) \circ g, D^1 g, \ldots, D^m g \in \mathbb{W}^{\sigma,\infty}.$$

By Lemma A.16, $D^m(u \circ g) \in \mathbb{W}^{\sigma,\infty}$. This implies $u \circ g \in \mathbb{W}^{s,\infty}$. □

Lemma A.13. *If $u \in \mathbb{W}^{M,p}$ for some integer M and $p \in [1, \infty]$ and $g \in \mathbb{K}^M$ and g is invertible. Then*

$$\|u \circ g\|_{M,p} \leq B_{M,p}(\|u\|_{M,p}, \|Dg\|_{M-1,p}, \|Dg^{-1}\|_{L^\infty}),$$

for some function $B_{M,p}$ that depends on M and p.

Proof. Since $g \in \mathbb{K}^M$, $Dg \in \mathbb{W}^{M-1,\infty}$. By definition,

$$\|u \circ g\|_{\mathbb{W}^{M,\infty}} = \sum_{\alpha : |\alpha| \leq M} \|D^\alpha (u \circ g)\|_{L^\infty}.$$

Since $g \in \mathbb{K}^M$ and g is invertible, Lemma A.11 implies $g^{-1} \in \mathbb{K}^M \in L^\infty$. For $p < \infty$,

$$\begin{aligned} \|(D^j u) \circ g\|_{L^p}^p &= \int_{\mathbb{R}^J} |(D^j u) \circ g(x)|^p dx \\ &= \int_{\mathbb{R}^J} |(D^j u)(x)|^p |\det Dg^{-1}(x)| dx \end{aligned}$$

$$\leq \|D^j u\|_{L^p}^p \cdot \sup_{x\in\mathbb{R}^J} |\det Dg^{-1}(x)|$$
$$\leq \|D^j u\|_{L^p}^p \|Dg^{-1}\|_{L^\infty}^J.$$

When $p = \infty$,

$$\|(D^j u)\circ g\|_{L^\infty} \leq \|D^j u\|_{L^\infty}.$$

For any $k \leq M$, each monomial in S_k defined by (A.25) involves at least one coordinate of $(D^j u)\circ g$. Note that $a \in L^p, b \in \mathbb{W}^{1,\infty}$ imply $\|ab\|_{L^p} \leq \|a\|_{L^p}\|b\|_{L^\infty}$. Letting $a = D^j u, b = g$, each monomial is in L^p. Therefore, we can find $\overline{S}_k$ such that

$$\left|S_k((D^1 u)\circ g,\ldots,(D^k u)\circ g, D^1 g,\ldots,D^k g)\right| \leq \overline{S}_k(\|u\|_{M,p}, \|Dg\|_{\mathbb{W}^{M-1,\infty}}, \|Dg^{-1}\|_{L^\infty}).$$

In particular, $\overline{S}_k \to 0$ when $\|u\|_{M,p} \to 0$. Thus,

$$\|u\circ g\|_{\mathbb{W}^{M,\infty}} \leq M^J \max_{0\leq k\leq M} \overline{S}_k(\|u\|_{\mathbb{W}^{M,\infty}}, \|Dg\|_{\mathbb{W}^{M-1,\infty}})$$
$$:= B_M(\|u\|_{\mathbb{W}^{M,\infty}}, \|Dg\|_{\mathbb{W}^{M-1,\infty}}, \|Dg^{-1}\|_{L^\infty}).$$

□

Lemma A.14 (Composition with a Lipschitz map preserves $\mathbb{W}^{\sigma,\infty}$ regularity)**.** *If $h \in \mathbb{W}^{\sigma,\infty}$ for some $\sigma \in [0,1)$ and $\phi : \mathbb{R}^J \to \mathbb{R}^J$ is Lipschitz with Lipschitz constant* Lip(ϕ)*, then $h\circ\phi \in \mathbb{W}^{\sigma,\infty}$ and*

$$[h\circ\phi]_{\mathbb{W}^{\sigma,\infty}} \leq [h]_{\mathbb{W}^{\sigma,\infty}} \operatorname{Lip}(\phi)^\sigma,$$

Proof. For $x \neq y$,

$$\|h(\phi(x)) - h(\phi(y))\| \leq [h]_{\mathbb{W}^{\sigma,\infty}} \|\phi(x)-\phi(y)\|^\sigma \leq [h]_{\mathbb{W}^{\sigma,\infty}} \operatorname{Lip}(\phi)^\sigma \|x-y\|^\sigma.$$

Taking the supremum over $x \neq y$ gives the claim. □

Lemma A.15 (Inversion preserves $\mathbb{W}^{\sigma,\infty}$ regularity)**.** *Let $M : \mathbb{R}^J \to \mathbb{S}_+^J$ be such that $M \in \mathbb{W}^{\sigma,\infty}$ for some $\sigma \in [0,1)$ and $\|M^{-1}\|_{L^\infty} \leq C$. Then $M^{-1} \in \mathbb{W}^{\sigma,\infty}$ and*

$$[M^{-1}]_{\mathbb{W}^{\sigma,\infty}} \leq C^2\,[M]_{\mathbb{W}^{\sigma,\infty}}.$$

Proof. Use the identity

$$M^{-1}(x) - M^{-1}(y) = M^{-1}(x)\big(M(y)-M(x)\big)M^{-1}(y).$$

Taking norms, dividing by $\|x-y\|^\sigma$, and taking suprema yields the estimate. □

Lemma A.16 ($\mathbb{W}^{\sigma,\infty}$ is closed in products)**.** *If $a, b \in \mathbb{W}^{\sigma,\infty} \cap L^\infty(\mathbb{R}^J)$ for some $\sigma \in [0,1)$, then $ab \in \mathbb{W}^{\sigma,\infty}$ and*

$$[ab]_{\mathbb{W}^{\sigma,\infty}} \leq \|a\|_{L^\infty} [b]_{\mathbb{W}^{\sigma,\infty}} + \|b\|_{L^\infty} [a]_{\mathbb{W}^{\sigma,\infty}}.$$

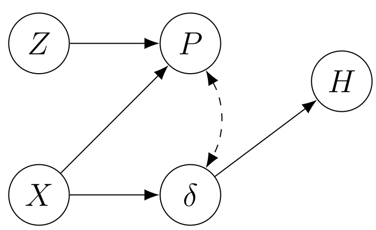
Proof. Write

$$a(x)b(x) - a(y)b(y) = a(x)(b(x) - b(y)) + b(y)(a(x) - a(y)).$$

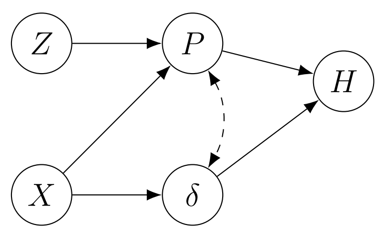
Divide by $\|x - y\|^\sigma$ and take suprema over $x \neq y$. □

Appendix B. Additional results and discussions not in the main text

B.1. **Faithfulness in causal graphs.** Given a candidate $H(\delta, P) = H$, consider the induced directed acyclic graph (DAG) of the variables (X, P, Z, δ, H). Depending on whether or not $H(\delta, P) = H(\delta)$, there are two possible DAGs shown in Figure B.1.[3] In the causal discovery literature (Spirtes et al., 2000), the joint distribution of $(X, P, Z, \delta, H) \sim Q_H$ is said to be *faithful* to a graph if, whenever two variables V_1 and V_2 are conditionally independent given some set of variables C, the correponding vertices in the DAG are *d-separated* by a set of vertices corresponding to C.



(a) H does not depend on P



(b) H does depend on P

Figure B.1. DAG for $(X, P, Z, \delta, H(P, \delta))$.

Suppose Q_H is known to be faithful to one of the two graphs, (a) or (b). Then, upon observing $Z \perp\!\!\!\perp H \mid X$, we can infer that Q_H is faithful to (a) and not (b), meaning H is constant in P. This is similar to the type of inference that Assumption 4 allows.

In this sense, Assumption 4 essentially requires that every Q_H is faithful to one of (a) and (b), depending on whether H varies with P—although Assumption 4 only speaks to whether the edge $P \to H$ is faithfully reflected by Q_H. Assumption 4 also strengthens the notion by requiring d-separation if H and Z are only *mean*

[3]Note that H is here a deterministic function of (δ, P) while each node in a DAG is typically assumed to have its own independent random variation.

independent given X, though our general identification argument accommodates the analogue of Assumption 4 with full independence instead.

B.2. **Identification in triangular models as faithfulness verification.** The faithfulness condition also connects to the nonparametric identification literature for triangular models. We detail this connection to Imbens and Newey (2009), Torgovitsky (2015), and D'Haultfœuille and Février (2015), whose arguments can be viewed as verifying faithfulness *without* needing an excluded proxy X but imposing instead that outcomes are scalar and monotone in δ.

Suppose $J = 1$ and let

$$S = \sigma(\delta, P), \quad P = f(Z, \omega).$$

Suppose that $\sigma(\cdot, P)$ is strictly increasing and $f(Z, \cdot)$ is strictly increasing (and their respective arguments are one-dimensional). Normalize the marginal distribution of ω to Unif$[0,1]$. Since ω is the conditional quantile of $P \mid Z$, we can assume that ω is observed. Lastly, assume that δ is continuously distributed and $\delta \mid \omega$ is also continuously distributed.

Under these restrictions, the identified set for σ is (Theorem 1, Torgovitsky (2015)):

$$\Theta_I = \left\{h^{-1}(\delta, p) \text{ strictly increasing in } \delta : (h(S, P), \omega) \perp\!\!\!\perp Z\right\}.$$

Analogous to Proposition 1, for identifying price counterfactuals, it suffices to show that, for h^{-1} strictly increasing in δ and $H(\delta, P) \equiv h(\sigma(\delta, P), P)$, the following analogue of faithfulness holds:

$$(H(\delta, P), \omega) \perp\!\!\!\perp Z \implies H(\delta, P) \text{ does not depend on } P. \tag{B.1}$$

Let $H = H(\delta, P)$. Imbens and Newey (2009) observe that

$$F_{H|P,\omega}(t \mid p, \omega) = \int \mathbf{1}(H(\delta, p) \le t)\, dF_{\delta|\omega}(\delta \mid \omega).$$

If the support of $\omega \mid P$ is $[0,1]$ for all P (Assumption 2, Imbens and Newey (2009)), we can then integrate the above with respect to the marginal distribution of ω:

$$\int_0^1 F_{H|P,\omega}(t \mid p, \omega) d\omega = \int \mathbf{1}(H(\delta, p) \le t)\, dF_\delta(\delta) = \mathbb{P}_{\delta \sim F_\delta}(H(\delta, p) \le t) = F_\delta\left(H^{-1}(t, p)\right). \tag{B.2}$$

Note that since $(H, \omega) \perp\!\!\!\perp Z$ and $P = f(Z, \omega)$, we have that

$$H \perp\!\!\!\perp P \mid \omega.$$

Thus $F_{H|P,\omega}(t \mid p,\omega) = F_{H|\omega}(t \mid \omega)$ does not depend on p. Thus, the left-hand side of Equation (B.2) does not depend on p. The only way for this to occur is if $H^{-1}(t,p)$ does not depend on p either, since F_δ is strictly increasing. This argument is in equation (5)–(7) of Imbens and Newey (2009) preceding their Theorem 3. We can thus view this argument as verifying the version of faithfulness (B.1).

Torgovitsky (2015) avoids the need in (B.2) to integrate over ω,[4] thus allowing for instruments with few support points.[5] Suppose that Z instead takes two values $\{z_0, z_1\}$. Let $\pi(p) = f(z_1, f^{-1}(p, z_0))$ such that for some identical value of ω,

$$\pi(p) = f(z_1, \omega), \quad p = f(z_0, \omega).$$

Torgovitsky (2015) observes that, since $Z \perp\!\!\!\perp (\delta, \omega)$,

$$\mathbb{P}(\delta \le H^{-1}(t, \pi(p)) \mid \omega) = F_{H|P,\omega}(t \mid \pi(p), \omega) = F_{H|P,\omega}(t \mid p, \omega) = \mathbb{P}(\delta \le H^{-1}(t,p) \mid \omega).$$

This implies that for any p, for all t,

$$H^{-1}(t, \pi(p)) = H^{-1}(t, p).$$

Torgovitsky (2015) imposes further assumptions that ensure that there is some value p_0 such that for any value p, the sequence $\pi(p), \pi(\pi(p)), \ldots$ approaches p_0.[6] Thus,

$$H^{-1}(t, \pi(p)) = H^{-1}(t, \pi(\pi(p))) = H^{-1}(t, \pi(\pi(\pi(p)))) = \cdots$$

equals $H^{-1}(t, p_0)$ given continuity. This ensures $H^{-1}(t,p) = H^{-1}(t, p_0)$, proving (B.1).

B.3. **Example of completeness without faithfulness.** We assume $J = 1$ for this example, though analogous constructions exist for $J > 1$. Suppose $Z, X > 0$ and

$$P \mid \delta, Z, X \sim \mathcal{N}(r(\delta, X)\tau(\delta, Z, X), \tau^2(\delta, X, Z)).$$

This construction is robust to taking monotone transformations of (δ, P), and thus we may view P as log price instead—for instance—if we would like price to be strictly positive. Let δ be supported on $[1, 2]$ with PDF

$$f(\delta \mid x) = \exp(r(\delta, x)^2/2)a(x)\mathbf{1}(\delta \in [1, 2])$$

[4]A similar argument appears in D'Haultfœuille and Février (2015).

[5]ω is the quantile of P in the $P \mid Z$ distribution. If there are only finitely many values for Z, then each price value has only finitely many quantiles that can be candidates for ω, meaning that $\omega \mid P$ cannot have full support.

[6]One example is to assume that p_0 is the left support end point of P and that $f(z_1, \omega) < f(z_0, \omega)$ for all $\omega > 0$ with $p_0 = f(z_0, 0) = f(z_1, 0)$. Then we have $\pi(p) < p$ approaching p_0.

for

$$a(x)^{-1} = \int_1^2 e^{r(\delta,x)^2/2}\, d\delta.$$

Proposition B.1. *Let $r(\delta, x) = x\delta$ and $\tau(\delta, z, x) = z$. Then the above construction specifies $(\delta, P) \mid (Z, X)$ for which Assumptions 2 and 3 hold but Assumption 4 fails.*

Proof. This choice is such that

$$\begin{aligned}\mathbb{E}\left[\mathbf{1}(P > 0) \mid \delta, Z, X\right] &= \mathbb{P}_{Q\sim\mathcal{N}(0,1)}\left(r(\delta, X)\tau(\delta, Z, X) + \tau(\delta, Z, X)Q > 0\right)\\ &= \Phi\left(r(\delta, X)\right)\end{aligned}$$

Thus

$$\mathbb{E}[\mathbf{1}(P > 0) \mid Z, X] = \mathbb{E}[\Phi\left(r(\delta, X)\right) \mid X] \equiv k(X).$$

Hence faithfulness does not hold. Assumption 3 holds by construction. Thus the remainder of the proof verifies that Assumption 2 holds.

The density of P is

$$\begin{aligned}f(p \mid x, \delta, z) &= \frac{1}{\sqrt{2\pi}} \exp\left(\log \frac{1}{\tau(\delta, Z, X)} - \frac{(p - r\tau)^2}{2\tau^2}\right)\\ &= \frac{1}{\sqrt{2\pi}} \exp\left(\log(1/\tau) - \frac{p^2}{2\tau^2} - \frac{r^2}{2} + \frac{yr}{\tau}\right)\end{aligned}$$

Under our choices for r, τ,

$$f(p \mid x, \delta, z) = \frac{1}{\sqrt{2\pi z}} \exp\left(-\frac{p^2}{2z^2} - \frac{r^2}{2} + p\delta\frac{x}{z}\right)$$

Thus

$$f(\delta, p \mid x, z) = \frac{1}{\sqrt{2\pi z}} a(x)\mathbf{1}(x \in [1, 2]) \exp\left(-p^2\frac{1}{2z^2} + p\delta\frac{x}{z}\right).$$

It is an exponential family supported on $[1, 2] \times \mathbb{R}$ with natural parameters

$$\eta(x, z) = \left(-\frac{1}{2z^2}, \frac{x}{z}\right)$$

It is possible to choose X, Z such that the support of $\eta(X, Z)$ contains an open set of $(-\infty, 0) \times (0, \infty)$. The sufficient statistics are

$$(T_1, T_2) = \left(P^2, \delta P\right)$$

Note that

$$\delta = |T_2/\sqrt{T_1}|, P = T_2/X$$

so that (δ, P) and (T_1, T_2) are bijective. Thus, take any function $h(\delta, P)$, we can write it in terms of the sufficient statistics $h(T_1, T_2)$. Since the support of $\eta(x, z)$ contains an open set in $(-\infty, 0) \times (0, \infty)$, T_1, T_2 are complete sufficient statistics. Thus

$$\mathbb{E}[H(\delta, P) \mid Z, X] = 0 \implies \mathbb{E}[\tilde{H}(T_1, T_2) \mid \eta] = 0 \implies \tilde{H} = H = 0.$$

□

B.4. **Identification with recentered instruments.**

Lemma B.1. *For $\mathcal{K} \subset L^2_F(X)$ and $\mathcal{H} \subset L^2_F(S, P)$,*

$$\begin{aligned}\Theta_I(\mathcal{H}, \mathcal{K}, F) &\subset \Theta_I(\mathcal{H}, L^2_F(X), F) \\ &= \left\{h \in \mathcal{H} : \mathbb{E}[h(S,P)\{R(X,Z) - \mathbb{E}[R(X,Z) \mid X]\}] = 0 \text{ for all } R \in L^2_F(X,Z)\right\}\end{aligned}$$

Proof. The subset inclusion is immediate. Given $h \in \Theta_I(\mathcal{H}, L^2_F(X), F)$ and any $R \in L^2_F(X, Z)$

$$\mathbb{E}[h\{R - \mathbb{E}[R \mid X]\}] = \mathbb{E}[k(X)\{R - \mathbb{E}[R \mid X]\}] = 0.$$

Conversely, take $R(X, Z) = \mathbb{E}[h(S, P) \mid X, Z]$. Then

$$\mathbb{E}[h(S,P)\{R(X,Z) - \mathbb{E}[R(X,Z) \mid X]\}] = \mathbb{E}\left[(\mathbb{E}[h(S,P) \mid X, Z] - \mathbb{E}[h(S,P) \mid X])^2\right] = 0.$$

This implies that

$$\mathbb{E}[h(S,P) \mid X, Z] = \mathbb{E}[h(S,P) \mid X] \equiv k(X) \in L^2_F(X).$$

□

B.5. **Bertrand–Nash pricing.** Here we show how Bertrand–Nash pricing implies Equation (13); see also Appendix A in Berry and Haile (2014). Consider the case of single-product firms.[7] As is well-known, the first-order condition for profit maximization of firm j with a constant marginal cost C_j is

$$\sigma_j(\delta, P) + (P_j - C_j)\frac{\partial \sigma_j(\delta, P)}{\partial P_j} = 0.$$

Provided the equilibrium is unique, the solution for the vector of P can therefore be written as a function of δ and C: i.e., Equation (13) holds.

[7]The argument extends to multi-product firms if the "ownership matrix" is conditioned upon in the same way characteristics $\tilde{X}$ are.