

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Orr, Scott

# Working Paper Response to Hong and Luparello (2024)

I4R Discussion Paper Series, No. 182

**Provided in Cooperation with:** The Institute for Replication (I4R)

*Suggested Citation:* Orr, Scott (2024) : Response to Hong and Luparello (2024), I4R Discussion Paper Series, No. 182, Institute for Replication (I4R), s.l.

This Version is available at: https://hdl.handle.net/10419/305230

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



# WWW.ECONSTOR.EU

# **INSTITUTE** for

No. 182 DISCUSSION PAPER SERIES

# **Response to Hong and Luparello** (2024)

Scott Orr

This paper responds to:

Hong, Joonkyo, and Davide Luparello. 2024. Robustness Report on "Within-Firm Productivity Dispersion: Estimates and Implications," by Scott Orr (2022). *IAR Discussion Paper Series* No. 181. Institute for Replication.

October 2024



# **I4R DISCUSSION PAPER SERIES**

I4R DP No. 182

# **Response to Hong and Luparello (2024)**

## Scott Orr<sup>1</sup>

<sup>1</sup>University of British Columbia, Vancouver/Canada

## OCTOBER 2024

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and metascientific work in the social sciences. Provided in cooperation with EconStor, a service of the <u>ZBW – Leibniz Information Centre for Economics</u>, and <u>RWI – Leibniz Institute for Economic Research</u>, I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

#### Editors

Abel Brodeur University of Ottawa Anna Dreber Stockholm School of Economics Jörg Ankel-Peters RWI – Leibniz Institute for Economic Research

E-Mail: joerg.peters@rwi-essen.de	Hohenzollernstraße 1-3	www.i4replication.org
RWI – Leibniz Institute for Economic Research	45128 Essen/Germany	

# Response to Hong and Luparello (2024)

## Scott $Orr^*$

September 30, 2024

I would like to thank Hong and Luparello (2024) for their effort in replicating, as well as partially extending, the results in Orr (2022). This replication report makes what I believe to be three key points:

- 1. The replication package of Orr (2022) is computationally reproducible (up to two minor caveats, which I very much appreciate the authors catching)
- 2. The demand estimates are sensitive to a 30% cost-share threshold pursued when constructing input price based IVs.
- 3. The demand estimation strategy does not work when applied to a different sample (the same industry, but a later time period).

In this brief note, I provide some minor commentary on each of these findings. I focus primarily on point 2, where my interpretation of their results differs slightly from their own. In particular, I believe these results largely show that the cost-share threshold does not matter much quantitatively — as long one is careful to choose thresholds that make economic sense — although I also acknowledge that the *precision* of the estimates can be sensitive to this threshold.

### 1) Computational Reproducibility

The fact that the paper replicates is, of course, good to hear. I am particularly happy to note that the authors were able to replicate the results from scratch, even after having to request the data themselves from MOSPI (I could not share the full data for replication as I signed a data confidentiality agreement when originally receiving the data). However, there are two minor caveats to computational reproducibility that I wish to comment on.

<sup>\*</sup>Sauder School of Business, University of British Columbia, scott.orr@sauder.ubc.ca

- i) The authors document that line 130 of asicc\_cleaning\_code.do will incorrectly drop industries 74 and 75 if you rely on the MOSPI provided ASICC09 codes csv file; this is because this file does not report a three digit ASICC code when there are no subdivisions within a given two digit code (this is true for industries 74 and 75). I thank the authors for catching this error. The source of the error is the following: when building the replication package, I relied on asicc2009\_mod.csv in line 129 of asicc\_cleaning\_code.do. This file is not the raw ASICC codes provided by MOSPI, but a file I constructed from the original file by manually inputting the relevant codes in the three digit code column of this file. I forgot to include this as a separate file in the replication package, thinking this was part of the MOSPI provided ASI data. To help future researchers on this front, I am providing the relevant asicc2009\_mod.csv to the I4R team as part of the replication package for this report. Inputting this file in the asicc\_2009 data folder will allow replicators to execute the full master file without error (alternatively, researchers can manually input the relevant 3-digit codes themselves, if they so desire).
- ii) While demand estimates perfectly replicate, the production function estimates of Hong and Luparello (2024) differ slightly from those in Orr (2022) to the second decimal. Hong and Luparello (2024) argue this due to numerical instability of STATA's matrix inversion when recovering marginal cost, which I believe is correct. The documentation for the particular command used for the marginal cost inversion luinv in STATA notes that different computers can give different results when a matrix is close to singular. While the particular matrix in question (an inverse of all cross-price derivatives in the market, multiplied by an ownership matrix) is not exactly singular, it appears to be close to singular due to improper scaling. In particular, all products (i.e. many thousands of different goods, which have very different sizes in terms of total revenue) are included in this matrix. I was not aware of this issue when I wrote the paper, and I appreciate the authors brining it to my attention. That said, the actual differences generated by these approximation errors end up being fairly small, and do not change the quantitative implications of the paper in any meaningful way, as the authors note.

#### 2) Cost Threshold Robustness

As a robustness exercise, the authors vary the cost-share threshold in Orr (2022) when constructing an instrument for output prices and market shares in demand estimation. This is a useful exercise. In this section, I provide some thoughts on how to properly interpret these results, emphasizing how the demand estimation quantities feedback into the paper's overall contribution — within-firm productivity estimates — which may not be obvious to all readers.

Orr (2022) relies on an input price based instrument, based on averages of prices paid by firms buying the same inputs used to produce machinery, equipment and parts (the industry of interest), but operating in other industries. The key idea is to leverage demand and supply shocks in other industries, which affect prices of inputs used in the machinery sector, and therefore their costs, prices, and market shares. The key threat to identification is that these average prices are actually driven by demand and supply shocks to the machinery sector (not other sectors), which would violate the exclusion restriction. To deal with this, Orr (2022) excludes average input prices (at the input product-code level), where machinery producers comprise more than 30% of total revenue observed going into purchases of that input product-code; the idea here being that if machinery producers comprise more than 30% of the market for that input, then average input prices are likely largely determined by machinery demand and supply shocks.

Since the case could be made for instead excluding products where machinery producers are say, 20% or 10% of the market (or potentially even less stringent thresholds, e.g. 40%, although it is important to note that these are likely to be more endogenous so perhaps less desirable), it is useful to see how the demand estimates change when varying this threshold. This is the key robustness exercise Hong and Luparello (2024) examine in their report.

Note that varying this threshold, however, involves tradeoffs. Making the threshold more stringent will make the instrument "more exogenous," (now average prices are only calculated for inputs where machinery is very unlikely to affect the overall level of prices) but at the cost of potentially reducing instrument strength through two margins; i) less observations (i.e., you end up dropping more product codes), and ii) you will end up relying on less important inputs for the machinery sector, since by construction machinery producers use these inputs less.

The key results of this exercise are summarized in Table 3 of the replication report. While varying the threshold does indeed affect the demand estimates, results do not change all that much unless you choose thresholds outside the 0.1 - 0.4 range. Note that larger thresholds (e.g. 0.5) imply that close to half of the market for a particular input is dominated by machinery producers; such input prices are unlikely to satisfy the relevant exclusion restriction. On the other hand, once we move below the 0.1 threshold, we're now forcing the instrument to rely on input prices that comprise a small proportion of the cost basket of machinery producers, which is very likely to negatively impact instrument strength. For these reasons, I believe thresholds outside the 0.1 - 0.4 range are unlikely to be effective

identification strategies.

Restricting attention to thresholds between 0.1 - 0.4, we can see from Table 3 that the point estimates for demand are quantitatively very similar to the original 0.3 threshold used in Orr (2022). Notably, both demand parameters change by at most 35% of the standard error of the original 0.3 threshold estimates. I would characterize this as fairly small changes, and largely leads me to conclude that threshold value is not first order important for determining demand parameters in this study.

The fact that the point estimates differ by less than half a standard error is important to bear in mind when trying to determine whether the final quantitative conclusions of the paper — notably, the implications of within-firm productivity heterogeneity — are likely to change when varying this threshold. In particular, a simple back-of-the envelope calculations indicates the block-bootstrapping procedure used in this paper already re-estimates the model for many demand parameters that are even more extreme than those reported in columns 2-4 of Table 3 of the replication report.

More precisely, the key final results reported in Section 5 of Orr(2022) on the implications of within-plant productivity heterogeneity quantify uncertainty over the demand parameters by parametrically bootstrapping the demand estimates, i.e. drawing these parameters from the multivariate standard normal implied by the asymptotic distribution of the IV estimator and re-estimating the model. Since the point estimates in columns 2-4 of Table 3 vary by less than 35% of each parameter's standard error, then, as a first pass (ignoring the covariance between the demand parameters, i.e. treating each parameter as drawn from a standard normal), we can calculate that around  $(1 - (\Phi(0.35) - \Phi(-0.35))) \times 100\% = 73\%$ block bootstrapped observations used to construct the final standard errors in Orr (2022) will likely involve more extreme (i.e. more different from the baseline point estimates), than the point estimates obtained by varying the cost-share threshold when constructing the demand instruments. Given this, if one relies on 95% confidence intervals to understand uncertainty in magnitudes for the final results in Section 5, these intervals are basically guaranteed to cover differences in final results generated by changing the cost-share threshold from 0.1 -0.4 many times over. As a result, the effect of these changes are unlikely to affect the final quantitative conclusions of the paper as their first-order effects are already contained within the confidence intervals of the block-bootstrapping procedure.

That said, it is worth recognizing that *precision* of the final estimates would almost certainly change with the cost-share threshold, given that instrument strength and standard errors fall when we rely on alternative IV thresholds. On this point I agree with Hong and Luparello (2024) that variability of the demand estimates is sensitive to the threshold rule. This means that relying on another threshold would likely increases the reported standard errors in Section 5. Readers concerned about the particular threshold used should be aware that the standard errors calculated in the second half of are likely a lower bound on the degree of demand uncertainty that might exist with lower or higher thresholds.

#### 3) Out of Sample Replication

The instrumental variable strategy for demand used in Orr (2022) does not work when applied to a later sample period (2011-2020). In particular, demand is found to be upward sloping, which makes it impossible to apply the key marginal cost inversion step required to recover within-firm productivity differences.

The key reason for this failure is documented in Table 5 of Hong and Luparello (2024); the input prices instruments are extremely weak in the later sample, with all first stage F-statistics falling below 2. Without strong, exogenous variation in input prices, an input price based identification strategy simply will not be able to estimate demand correctly.

I agree with the authors' suggestion at the end of the paper that determining alternative IVs to identify demand would be a valuable path for future research. While this is an old problem, it is far from a settled problem. Many classes of instruments haven been proposed by a number of authors; see the recent review articles in the Handbook of Industrial Organization by Berry and Haile (2021) and Gandhi and Nevo (2021), both of which provide careful discussions of appropriate IVs for demand estimation. While input price instruments are simply one possible way forward — which appear to have some explanatory power in the 2000-2007 sample, but not so in later samples — many alternative sources of variation exist for identifying demand. Which particular strategy should work in a particular context is hard to predict *ex ante*. I would encourage researchers to adapt their demand estimation routine to the particular context they have in mind when applying the methodology developed in Orr (2022) for recovering within-firm productivity differences.

Note, however, that having continued to work on related problems since the publication of Orr (2022), there is one candidate of instrument that I have found to work well in other contexts that may be worth mentioning in the interest of stimulating further researcher. More precisely, relying on productivity shocks (i.e. TFP — the core objects of interest in Orr 2022) as an instrument for demand has some desirable properties; it is naturally "excluded" from demand (consumers do not care about productivity except through prices), and tends to be strongly correlated with prices. This strategy was used quite successfully in Foster et al. (2008) albeit in a setting where quality differentiation was less important, so concerns that TFP would be correlated with unobserved quality shifters in the demand function could largely be avoided.

Recently, Orr and Tabari (2024) implement a similar strategy, using the Indian Prowess

data from 1989 – 2019. To account for quality shifters being correlated TFP (a potential first order concern documented by Orr 2022, as well as many others, e.g. Jaumandreu and Yin 2020, Forlani et al. 2023, Caselli et al. 2024), they include a series of product code fixed effects (to control for product-specific quality characteristics that are fixed over time), as well as firm-year fixed effects (to control for quality-upgrading at the firm-level over time). This strategy appears to produce fairly reasonable demand elasticity estimates, and always has a very strong first stage.

It would be interesting to see if this type of estimation strategy can be implemented in tandem with the Orr (2022) approach to recovering productivity. Actually implementing this approach would require researchers tackle another interesting problem; if the demand system allows for firm-product varying markups, then recovering TFP by product requires estimates of the demand system. The challenge then becomes; how can we get estimates of TFP to help us identify demand, if we *need* demand to recover TFP! I do not think this circularity problem is impossible to solve. In fact, I suspect there exists an iterative procedure, or perhaps a system of simultaneous equations estimation approach, that could be used to operationalize this strategy. Working out the theory behind when such an estimator would be feasible, and implementing it in a wider class of samples, would, I believe, be a very interesting paper.

While the above is surely speculative, it is provided in the spirit of Hong and Luparello (2024) suggestion that identifying more reliable IVs for demand estimation would be a valuable path for future research, which I wholeheartedly agree with.

### Conclusion

In closing, I appreciate the replication of my work presented by Hong and Luparello (2024), and hope my discussion of their report helps other readers fully appreciate their findings, as well as potentially stimulate further research on the broad topic at hand.

## References

- Berry, Steven T and Philip A Haile, "Foundations of demand estimation," in "Handbook of industrial organization," Vol. 4, Elsevier, 2021, pp. 1–62.
- Caselli, Mauro, Arpita Chatterjee, and Shengyu Li, "Productivity and Quality of Multi-product Firms," *Working Paper*, 2024.
- Forlani, Emanuele, Ralf Martin, Giordano Mion, and Mirabelle Muûls, "Unraveling firms: Demand, productivity and markups heterogeneity," *The Economic Journal*, 2023, 133 (654), 2251–2302.

- Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?," *American Economic Review*, 2008, *98* (1), 394–425.
- Gandhi, Amit and Aviv Nevo, "Empirical models of demand and supply in differentiated products industries," in "Handbook of industrial organization," Vol. 4, Elsevier, 2021, pp. 63–139.
- Hong, Joonkyo and Davide Luparello, "Robustness Report on "Within-Firm Productivity Dispersion: Estimates and Implications," by Scott Orr (2022)," *Working Paper*, 2024.
- Jaumandreu, Jordi and Heng Yin, "Cost and product advantages: A firm-level model for the chinese exports and industry growth," *Working Paper*, 2020.
- **Orr, Scott**, "Within-Firm Productivity Dispersion: Estimates and Implications," *Journal of Political Economy*, 2022, 130 (11), 2771–2828.
- **and Mokhtar Tabari**, "Decomposing the Within-Firm Productivity Gains From Trade: Evidence from India," *Working Paper*, 2024.