Have We Got News for You: Firm-Level Evidence on the Optimal Choice of Expected Capacity Utilization

Niklas Amberg
Sveriges Riksbank

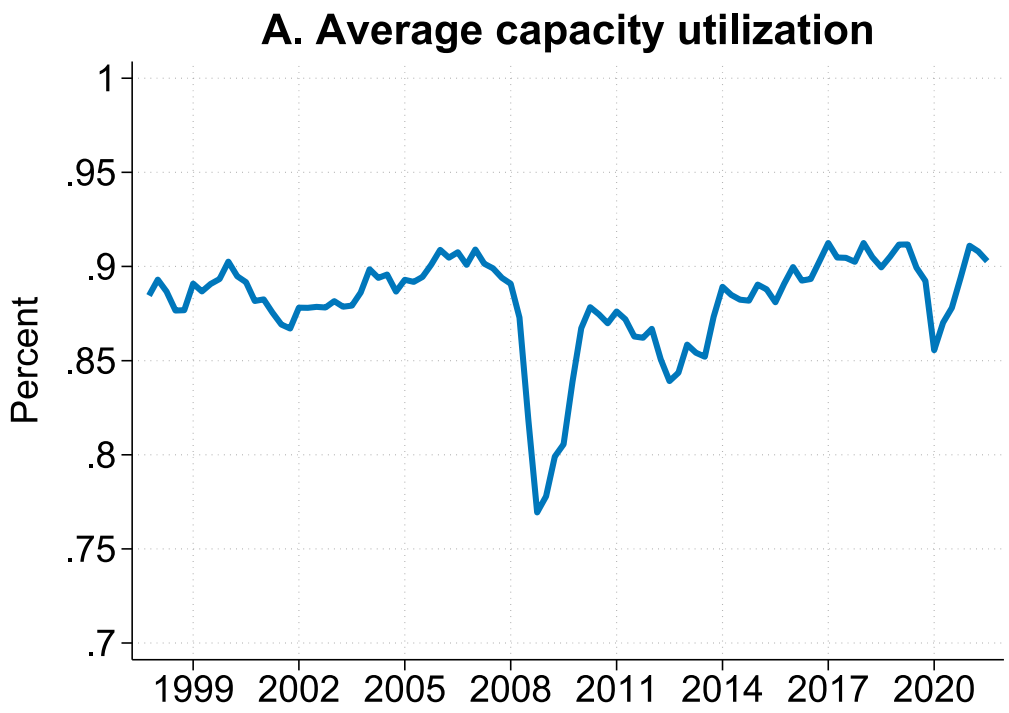
Richard Friberg
Stockholm School of Economics and CEPR

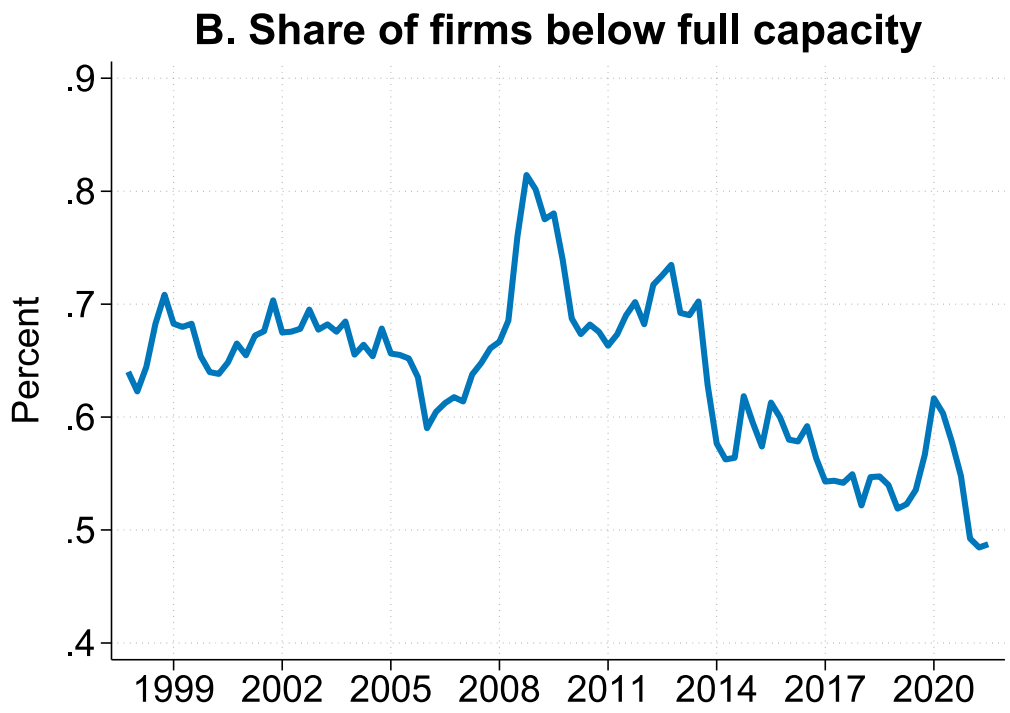
Chad Syverson
University of Chicago Booth School of Business and NBER

EEA Congress 2025 (Bordeaux School of Economics) – August 25, 2025

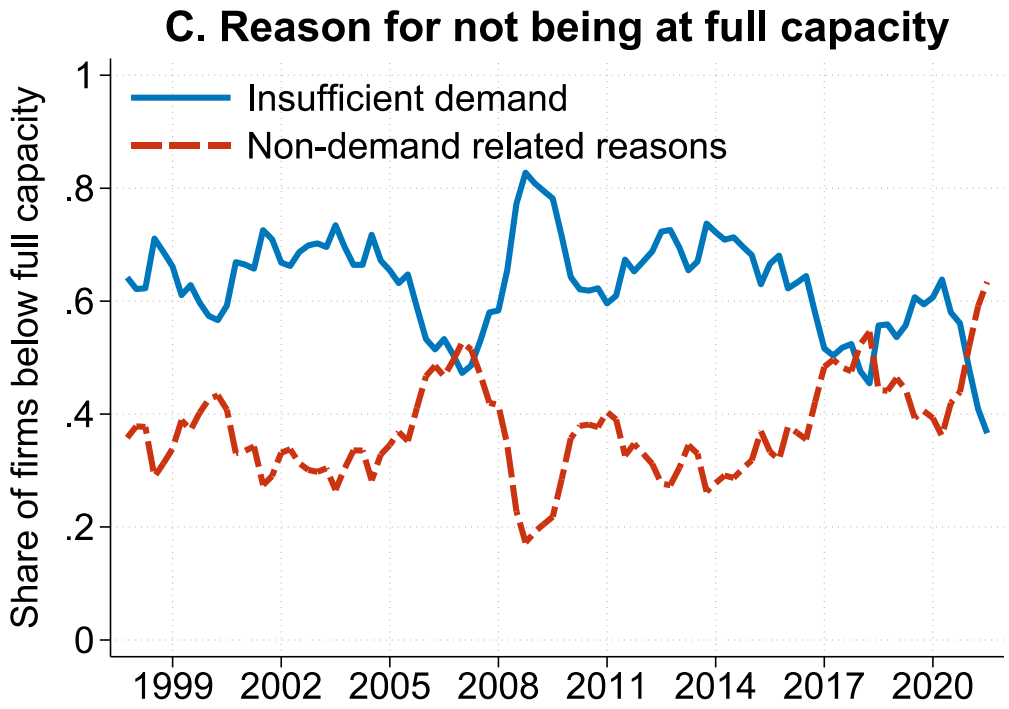
The opinions expressed in this presentation are the sole responsibility of the presenter and should not be interpreted as reflecting the views of Sveriges Riksbank

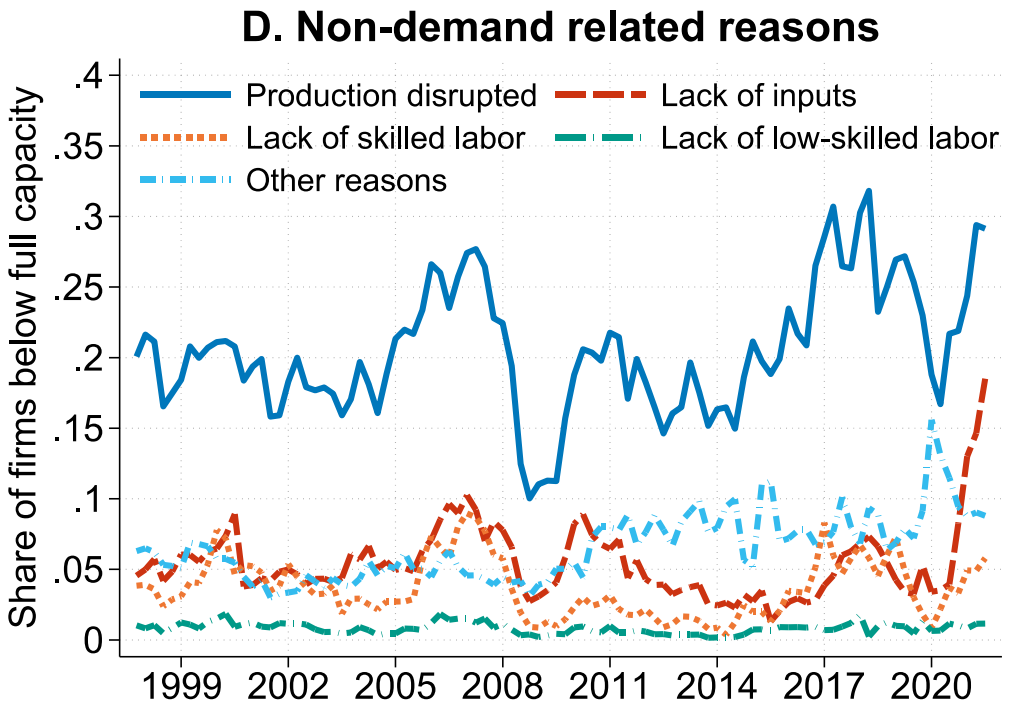
Fact 1: Most firms routinely operate below full capacity, even outside of recessions



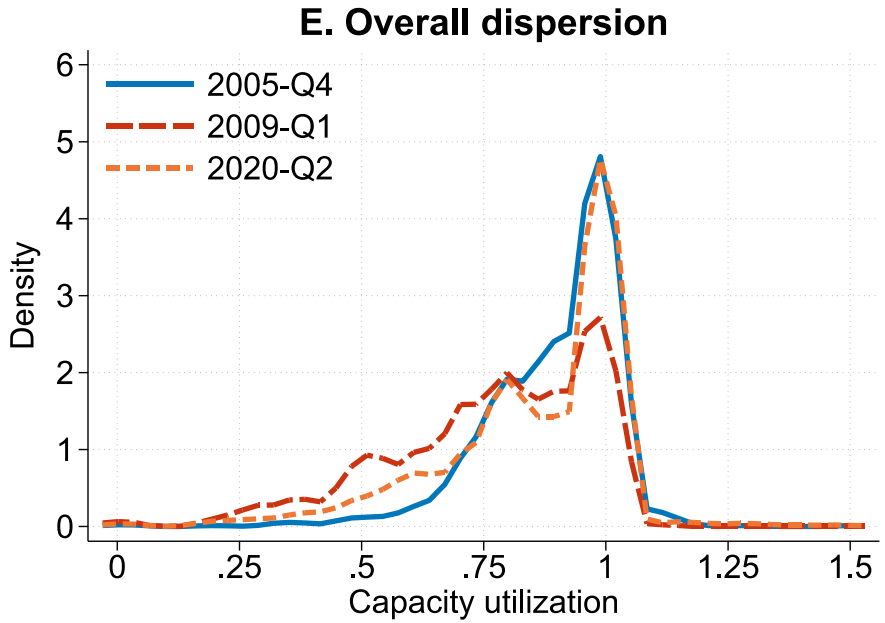


Fact 2: “Insufficient demand” is the main reason for operating below full capacity





Fact 3: The variation in capacity utilization across firms and over time is mainly idiosyncratic



	Dependent variable: Capacity utilization ($CU_{i,t}$)			
	(1)	(2)	(3)	(4)
	Time FE	Industry-time FE	Firm FE	Industry-time and firm FE
R^2	0.033	0.089	0.433	0.488
Number of obs.	81,272	81,272	81,272	81,272

This paper

Goal: Figure out why firms often operate well below full capacity, and why some firms are more prone to do so than others

Hypothesis: Firms optimally operate with idle capacity when they face uncertainty about future demand

- That is, idle capacity serves as capacity buffers that firms rationally maintain to manage uncertainty in demand

Approach:

1. Work out testable hypotheses about the effect of demand uncertainty on expected capacity utilization using a simple newsvendor model
2. Test the predictions empirically using rich Swedish micro data

A simple newsvendor model of capacity utilization

A simple newsvendor model

We use a model of the newsvendor type (Arrow, Harris and Marschak, 1951)

- Traditionally formulated for inventory, but its logic also applies to capacity decisions

Ours is the simplest possible version and has the following key features:

- A firm sets capacity K with a marginal cost c before demand is known. There are no production costs other than the cost of capacity.
- Demand q is drawn from the probability density function $f(q)$, known to the firm
- The firm sells at the exogenously given price p

The firm's problem is to set capacity to maximize expected profits:

$$\max_K \int_0^K (pq - cK) f(q) dq + \int_K^\infty (pK - cK) f(q) dq$$

Optimal capacity

The first-order condition can be expressed as

$$c = p\big(1 - F(K^*)\big)$$

The optimal capacity level is thus the point at which the marginal cost of capacity equals the expected marginal revenue of capacity

Denoting the inverse CDF by F^{-1} and solving for capacity yields

$$K^* = F^{-1}\left(\frac{p - c}{p}\right)$$

Optimal capacity K^* is thus a quantile of the demand distribution, determined by the markup (expressed in the Lerner index form)

The effects of changes in demand uncertainty

Define expected capacity utilization as:

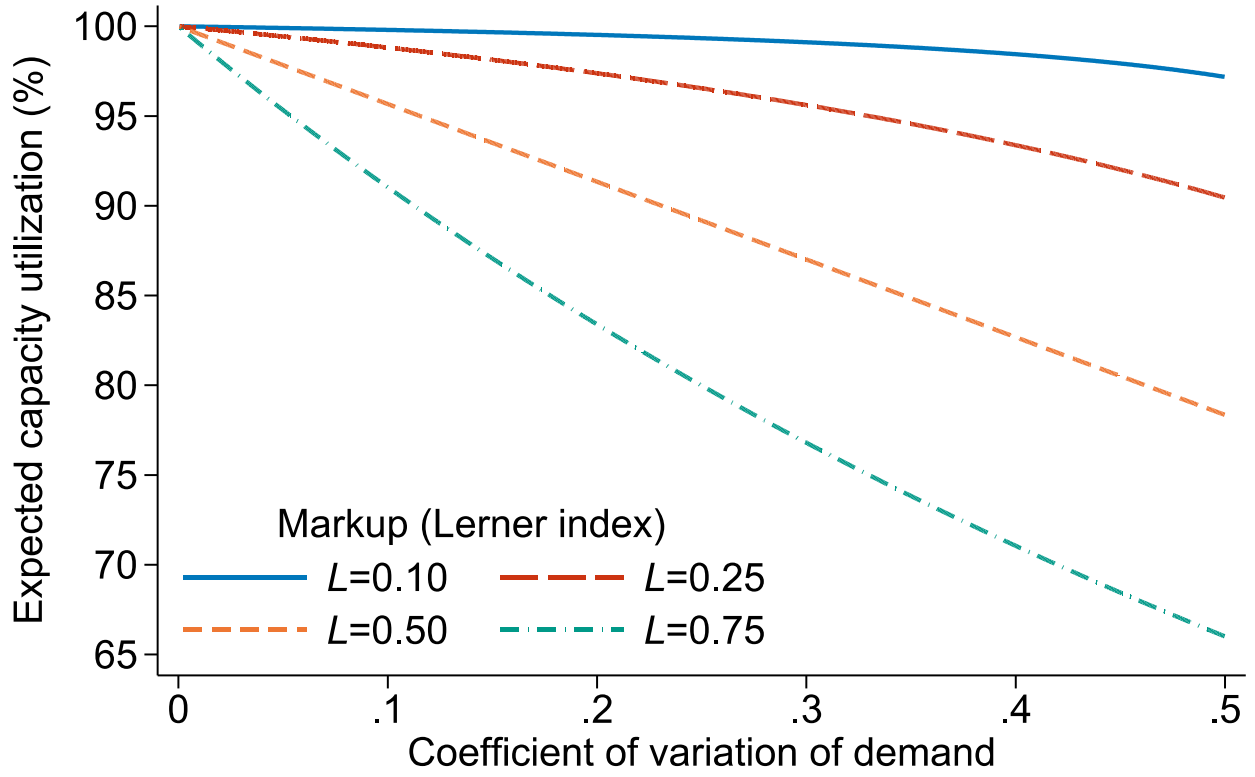
$$\frac{\bar{q}}{K^*} = \frac{\int_0^{K^*} qf(q)dq + (1 - F(K^*))K^*}{K^*}$$

where $\bar{q}$ is expected quantity sold (notice that $\bar{q} < E(q)$)

Butters (2019) proves that if uncertainty is dispersively ordered, then optimal expected capacity utilization decreases as uncertainty goes up

- G is a dispersive order of F if every possible interquantile range is wider apart for G than for F (that is, if $F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha)$ for any $0 < \alpha < \beta < 1$)

Simulated capacity utilization as a function of demand uncertainty and markups for uniform demand ($F \sim U[a, b]$)



Two testable predictions

1. Higher demand uncertainty is associated with lower capacity utilization
2. The effect of demand uncertainty on capacity utilization is stronger the higher is a firm's markup

Taking the predictions to the data

Data and sample

Main data source: Konjunkturbarometern [*The Business Cycle Barometer]*

- The largest business survey in Sweden, conducted by a government agency. Part of the "*Joint Harmonized EU Programme of Business and Consumer Surveys*."
- Monthly or quarterly frequency, depending on question. Response rate: ~50 percent.
- Firms with 100+ employees are always in the survey; smaller firms are randomly sampled. A new sample of about 5,800 firms is drawn once per year.
- Key survey questions for our purposes: Capacity utilization and demand uncertainty

Sample: Manufacturing firms appearing in the data during 2021Q2–2024Q1

- Around 7,400 observations across 1,000 firms
- Typical firm is old, mid-sized company, but lots of dispersion in both age and size

Empirical specification

Our baseline empirical model is:

$$CU_{it} = \alpha_i + \psi_{jt} + \sum_{k=2}^{4} \beta_k \cdot 1\{DU_{it} = k\} + \boldsymbol{\Omega} \cdot \boldsymbol{X}_{it} + \varepsilon_{it}$$

CU_{it}: Firm i's capacity utilization in quarter t

DU_{it}: Firm i's perceived, forward-looking demand uncertainty in quarter t

Demand uncertainty is the categorical answer to the survey question:

"The future development of our business situation is currently [easy to predict / fairly easy to predict / fairly difficult to predict / difficult to predict]."

Empirical specification: Comments

The model makes predictions about *expected* capacity utilization, but what we observe in the data is *realized* capacity utilization: $CU_{it} = \overline{CU}_{it} + \epsilon_{it}$

- Thus: critical to control for output shocks realized after capacity decisions (ϵ_{it})
- Our control is the categorical answer to the question: "Our production volume has in the past three months [increased / not changed / decreased]"

It is also important to separate uncertainty from predictable future variation

- We do so by including the answer to the question "We expect our production volume in the coming three months to [increase / not change / decrease]" as a control

Some specifications also include controls for inventory levels

Table 2: The effect of demand uncertainty on capacity utilization

	Dependent variable: Capacity utilization ($CU_{i,t}$)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicting the future development of our business situation is currently... [Omitted category: Easy]							
–Fairly easy	–0.015	–0.018	–0.020	–0.018	–0.018	–0.019	0.008
	(0.016)	(0.015)	(0.015)	(0.015)	(0.011)	(0.019)	(0.012)
–Fairly difficult	–0.061***	–0.050***	–0.053***	–0.049***	–0.043***	–0.050**	–0.001
	(0.016)	(0.015)	(0.015)	(0.015)	(0.012)	(0.022)	(0.012)
–Difficult	–0.095***	–0.070***	–0.072***	–0.069***	–0.065***	–0.069***	–0.016
	(0.018)	(0.017)	(0.017)	(0.017)	(0.013)	(0.022)	(0.014)
Our production volume has over the past three months... [Omitted category: Not changed]							
–Increased		0.029***	0.034***	0.032***	0.027***	0.032***	0.018***
		(0.005)	(0.005)	(0.005)	(0.003)	(0.006)	(0.004)
–Decreased		–0.098***	–0.094***	–0.088***	–0.068***	–0.069***	–0.057***
		(0.007)	(0.007)	(0.007)	(0.004)	(0.006)	(0.008)
We expect our production volume over the next three months to... [Omitted category: Remain the same]							
–Increase			–0.021***	–0.021***	–0.010***	–0.018***	–0.007**
			(0.005)	(0.005)	(0.003)	(0.006)	(0.004)
–Decrease			–0.017**	–0.014**	–0.002	–0.001	–0.008
			(0.007)	(0.006)	(0.004)	(0.005)	(0.007)
Industry × time FE	Yes	Yes	Yes	Yes	No	No	No
Firm and time FE	No	No	No	No	Yes	Yes	Yes
Inventory controls	No	No	No	Yes	Yes	Yes	Yes
Number of obs.	6,603	6,603	6,603	6,603	6,603	3,387	3,332
Number of firms	953	953	953	953	953	695	710
R^2	0.294	0.360	0.363	0.375	0.701	0.777	0.733

Mean idle capacity in the sample: **17%**

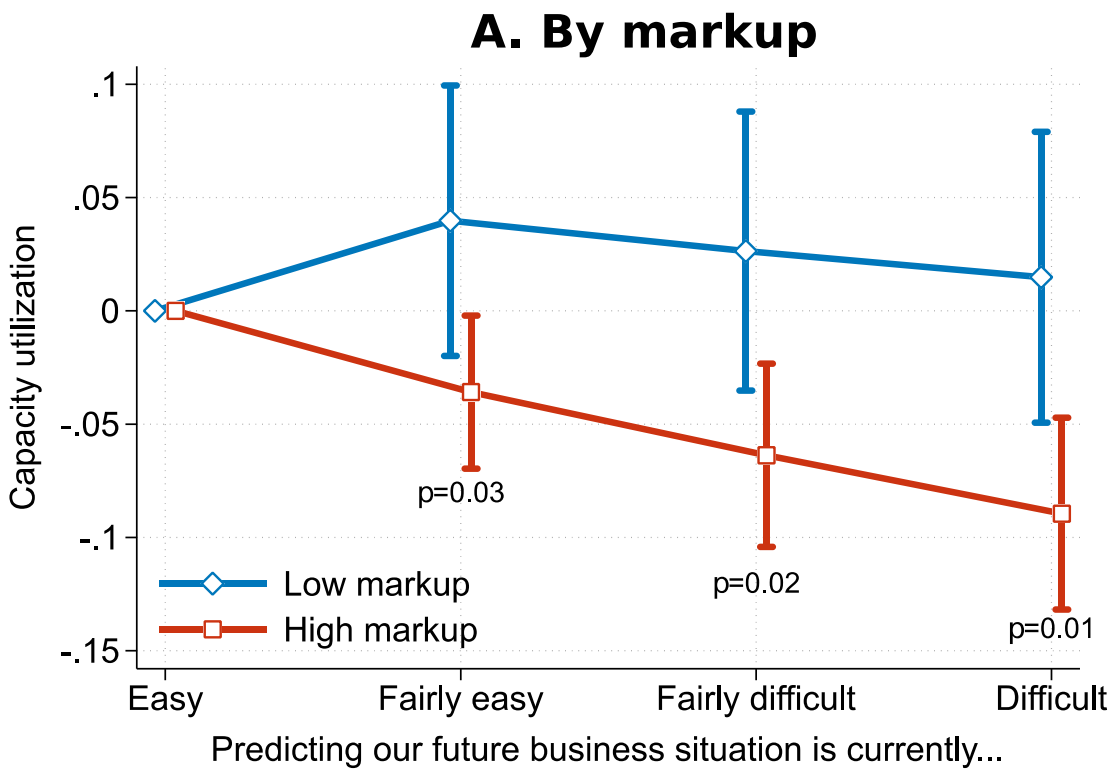
The role of markups: Empirical approach

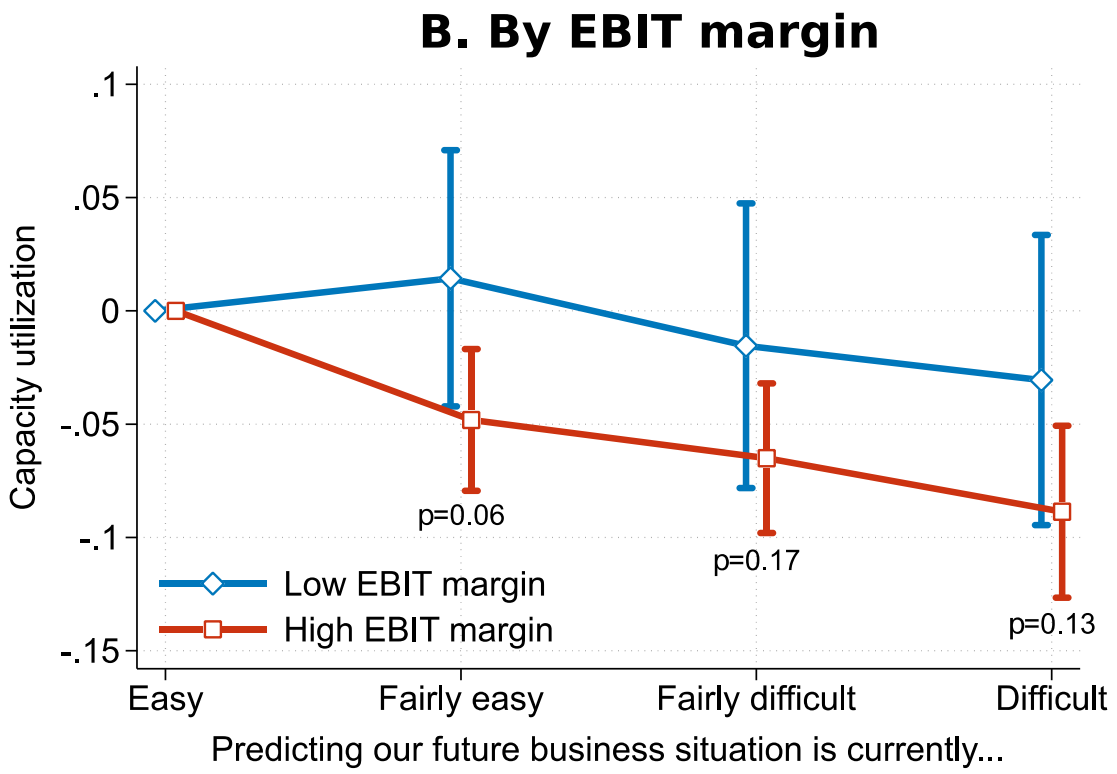
Second prediction from the newsvendor model: The effect of demand uncertainty on capacity utilization is stronger the higher is a firm's markup

We proceed as follows to test this prediction:

1. Compute two different markups using financial-accounts data: De Loecker and Warzynski (2012) and EBIT/Sales. Take averages over 2016–2020.
2. For each markup measure: Classify firms in the top tercile as high-markup firms and firms in the bottom tercile as low-markup firms
3. Estimate the baseline specification separately for high- and low-markup firms

The role of markups: Results





Conclusion and discussion

We show that uncertainty and markups are important drivers of firms' capacity decisions, in line with the predictions of the newsvendor model

- These findings complement the existing capacity literature, which largely focuses on the macroeconomic consequences of capacity constraints and idle capacity

Some concluding thoughts

- Excess capacity is a type of insurance and should be evaluated as such
- Our results point to a possible additional distortion of market power: higher costs of maintaining unused capacity
- Capacity utilization has trended down in the U.S. in recent decades – our findings point to increases in uncertainty and/or markups as possible drivers of this trend

Extras

Data source: Statistics Sweden

We establish the stylized facts using the micro data underlying the official Swedish industrial capacity-utilization series (from Statistics Sweden)

- Survey data covering firms in mining and manufacturing sectors with ≥50 employees
- Data is of quarterly frequency and spans the period 1998Q1–2023Q1
- About 91,000 observations across 3,000 firms (response rate ~80 percent)
- Key questions: capacity utilization and reasons for operating below full capacity

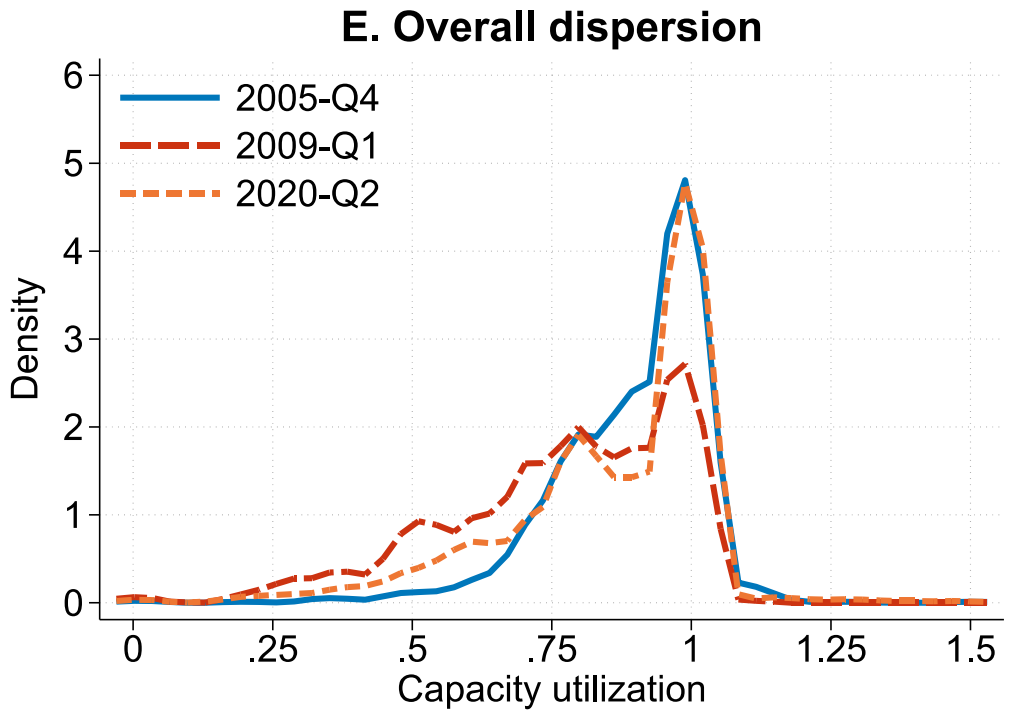
Exact wording of capacity-utilization question (translated to English by ChatGPT)

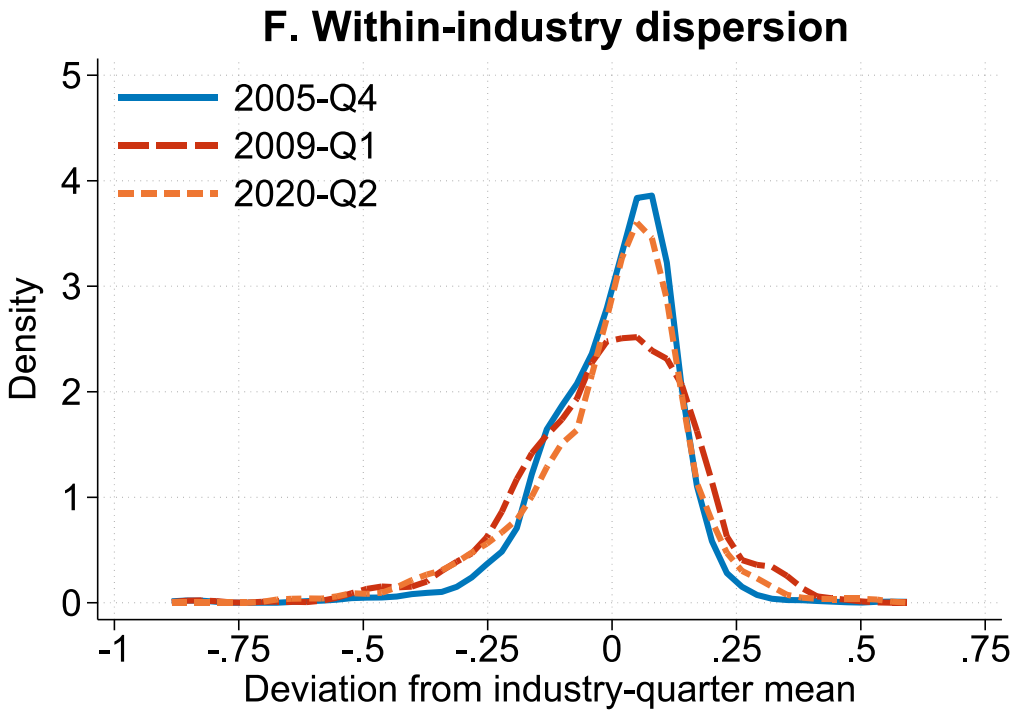
Capacity utilization is defined as the ratio between actual production and full production capacity. Actual production refers to the extent to which the industrial machinery is utilized under the prevailing production method during the quarter. Full production capacity refers to the level of production that can be achieved with the existing machinery and prevailing production method during the quarter.

Note that:

- Variations in production capacity due to seasonal factors, such as holidays, should not be included.
- Capacity utilization can exceed 100 percent, for example, in cases of overtime or when additional shifts are introduced.
- Use the working hours and shift patterns that can be considered normal.
- If measures have been taken to change production capacity, the new situation should be considered normal.

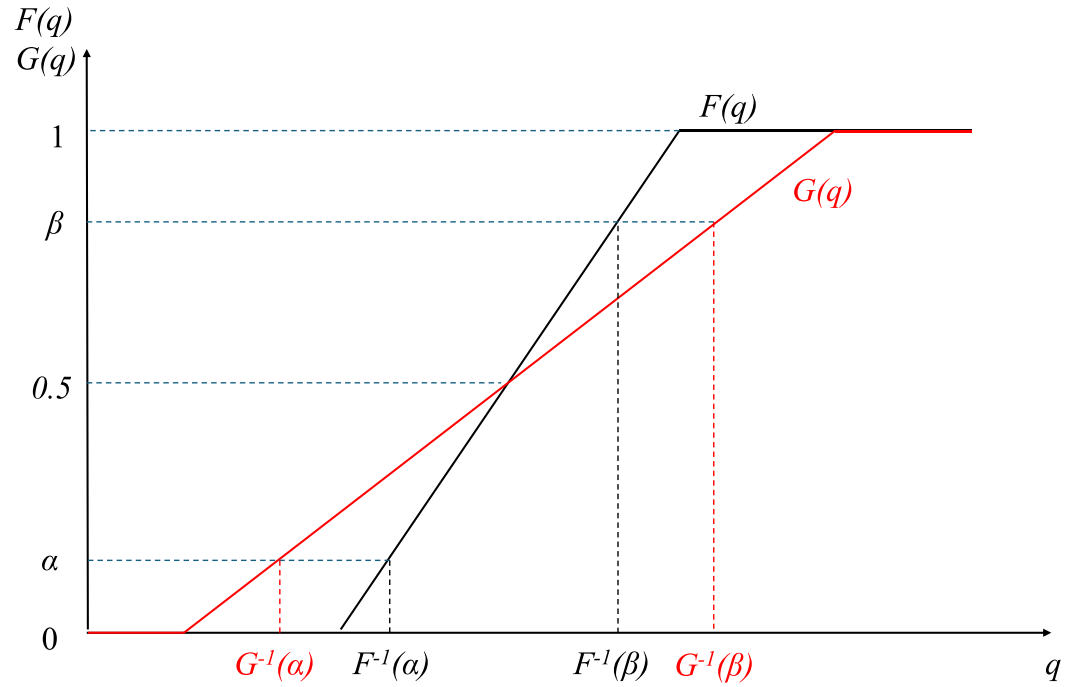
Fact 3: Variation in capacity utilization across firms and over time is mainly idiosyncratic





Dispersive orders: Definition and an illustration

Definition: *G is a dispersive order of F if $F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha)$ for any $0 < \alpha < \beta < 1$, i.e., if every interquantile range is wider for G than for F.*



Demand uncertainty and expected capacity utilization

Proposition 1 (Butters, 2019): *Consider two random variables with CDFs F and G that are both continuous and symmetric and have the same mean and median. If G is riskier than F in the sense that G is a dispersive order of F, then expected capacity utilization is lower under G than under F:*

$$\frac{\bar{q}_G}{K_G^*} < \frac{\bar{q}_F}{K_F^*}$$

Summary statistics for key variables