Microdata-Based Output Gap Estimation and Business Cycle Analysis Using Enterprise Surveys: Evidence from Poland

Mariusz Górajski* Magdalena Ulrichs*

September 2026

Abstract

This paper develops a quarterly, firm-level, production-function-based measure of the output gap using microdata from Statistics Poland's Quarterly Non-Financial Enterprises Survey (QNFES) and Business Tendency Survey (BTS) over 2008Q1–2025Q1, extending to quarterly frequency the annual firm-level framework of Błażej et al. (2025). We estimate sector-specific Cobb–Douglas production functions separately for manufacturing, construction, and services, augmented with a direct survey-based measure of capacity utilisation, applying the Olley–Pakes control-function approach with the Ackerberg–Caves–Frazer correction to firm-level panel data; a Levinsohn–Petrin/ACF specification is estimated as a robustness check and yields closely comparable output-gap paths. Firm-level inputs and output are seasonally adjusted and filtered to extract their potential components, which are then substituted into the estimated production functions to obtain firm-level potential output. Aggregation across firms yields economy-wide and sectoral quarterly output gaps, as well as a decomposition into capital, labour, and capacity-utilisation components.

The resulting microdata-based output gap, defined as the log difference between aggregated actual and potential output, closely tracks fluctuations in the Polish economy and identifies cyclical turning points at quarterly frequency. We use this measure to construct business-cycle indicators — including business cycle clock diagrams — that identify the cyclical phase of the economy and the duration of expansions and contractions in Poland over 2008–2025. Relative to the annual framework, the quarterly setting additionally requires firm-level seasonal adjustment and a boosted Hodrick–Prescott filter for trend extraction.

We further assess the macroeconomic relevance of the proposed measure using GMM estimates of Phillips-curve specifications. The microdata-based output gap is a statistically significant and economically relevant predictor of consumer price inflation. We then evaluate the real-time reliability of the sectoral estimates directly, using an expanding-sample vintage exercise that compares nine successive estimation vintages, the largest of which serves as a pseudo-final benchmark – an assessment rarely undertaken for microdata-based output-gap measures. This exercise reveals substantial heterogeneity across sectors: the manufacturing output gap shows no statistically significant end-point bias, whereas the construction

*University of Łódź, Department of Econometrics; Statistics Poland, Survey Methodology Department.

and, particularly, services gaps are subject to statistically significant real-time revisions. For services, the root-mean-square real-time revision exceeds the standard deviation of the estimated gap itself.

Overall, the results show that firm-level survey data provide a timely and granular measure of the cyclical position of the Polish economy at quarterly frequency, across sectors well beyond manufacturing. The proposed output gap is informative for business-cycle monitoring and inflation analysis, although substantial cross-sectoral differences in real-time revisions point to important limitations in the reliability of sector-specific estimates.

JEL classification: C33, C82, D25, E32, E37

Keywords: potential output; output gap; business cycle; production function estimation; Olley–Pakes method; business tendency survey; quarterly microdata

1 Introduction

The output gap – the deviation of actual from potential output – remains one of the most closely watched, yet also most contested, indicators in applied macroeconomics (see Orphanides and van Norden, 2002; Basu and Fernald, 2009; Havik et al., 2014; Blondeau et al., 2021). It plays a central role in assessing inflationary pressures through the Phillips curve, informs the calibration of monetary policy rules, and provides a key input into the calculation of the cyclically adjusted (structural) fiscal balance used in EU fiscal surveillance. Because potential output is unobservable, however, its measurement necessarily depends on modelling choices, and resulting estimates can be particularly unstable at the end of the sample – the so-called "real-time" or "end-point" problem (Orphanides and van Norden, 2002).

Existing production-function methodologies for output-gap estimation – including those used by the OECD, the European Commission (EC), and the US Congressional Budget Office (CBO) – typically apply a Cobb–Douglas or CES production function to *aggregate*, macro-level time series of capital, labour, and, in some cases, capacity utilisation (see Havik et al., 2014; Chalaux and Guillemette, 2019; Blondeau et al., 2021). Such aggregate approaches mask substantial heterogeneity across firms, sectors, and regions and typically do not incorporate a direct, observable measure of capacity utilisation, instead relying on indirect proxies or omitting utilisation altogether. A small but growing literature has explored more disaggregated approaches using firm-level or sectoral data to construct output-gap or capacity-utilisation-gap measures. For instance, a capacity-utilisation gap for Turkey has been constructed from firm-level survey data, while sectoral output gaps for Bulgaria have been estimated by applying a Hodrick–Prescott filter to sectoral gross value added (see Havik et al., 2014). Building on this line of research, Błażej et al. (2025) propose an *annual* firm-level, production-function-based output gap for Poland, estimated on firm-level panel data augmented with a survey-based measure

of capacity utilisation. That measure correlates strongly with OECD and EC output-gap estimates while providing greater explanatory power for inflation in Phillips-curve regressions than either the institutional estimates or a univariate filter applied to aggregate GDP; Błażej et al. (2025b) use the same annual framework to study the synchronisation of regional and sectoral gaps within Polish manufacturing.

This paper is a direct successor to Błażej et al. (2025): it takes that annual framework as its point of departure and rebuilds it at quarterly frequency. The higher-frequency perspective matters because both economic activity and inflation dynamics evolve at sub-annual frequencies, and indicators such as the cyclically adjusted budget balance are conventionally defined at quarterly or higher frequency, so an annual output-gap series cannot, by construction, resolve within-year cyclical turning points. We construct the quarterly output gap using Statistics Poland's Quarterly Non-Financial Enterprises Survey (QNFES, firms with 50 or more employees) matched with the quarterly BTS capacity-utilisation module, covering 2008Q1–2025Q1.

The move from annual to quarterly firm-level data is not a mechanical change of frequency, and it motivates three departures from the annual framework. First, firm-level flow and capacity-utilisation variables must be seasonally adjusted before trends are extracted, since quarterly firm-level series exhibit pronounced and heterogeneous seasonal patterns that are absent from annual data by construction; we adjust each firm's series individually using an automatically selected ETS model. Second, trend extraction itself becomes more demanding, because a quarterly panel carries substantially more high-frequency variation and because end-of-sample trend uncertainty is the binding constraint on real-time usefulness; we therefore replace the conventional Hodrick–Prescott filter of the annual framework with the boosted HP filter of Phillips and Shi (2021), whose data-driven iteration is designed to remove the residual low-frequency variation that a single HP pass leaves in the cyclical component. Third, the quarterly resolution makes formal business-cycle dating meaningful: we date peaks and troughs on a band-pass-filtered gap, measure the duration of expansions and contractions in quarters, and summarise the economy's position with "business cycle clock" diagrams (European Commission, 2023) of the kind published by statistical agencies, which locate the economy in the expansion, slowdown, recession, or recovery quadrant of the phase plane. None of these three elements has a counterpart in the annual framework. Beyond them, we extend the coverage from manufacturing to construction and services, add a real-time vintage assessment, and check the sensitivity of the results to the choice of control-function proxy.

Our contribution is fivefold. First, we provide, to the best of our knowledge, the first quarterly, firm-level, production-function-based output gap for a European economy – the quarterly counterpart of the annual measure of Błażej et al. (2025) –, covering Poland over 2008Q1–2025Q1 and encompassing several major cyclical episodes, including the Global Financial Crisis, the 2012–2013 slowdown, the 2020 COVID-19 shock, and the economic

disruptions following the 2022 full-scale invasion of Ukraine by Russia. The underlying estimation exercise is itself of independent interest: we estimate 62 division-level control-function production functions on quarterly firm-level data in which capacity utilisation enters as a *directly observed* input, so that measured total factor productivity can be separated into a utilisation-adjusted component and a utilisation component rather than being left to absorb both. Second, we show how the quarterly microdata-based output gap can be decomposed into capital-, labour-, and capacity-utilisation-based components and into sectoral contributions from manufacturing, construction, and services – the latter two sectors receiving comparatively little attention in the existing production-function output-gap literature, which has focused predominantly on manufacturing. The decomposition reveals substantial variation in the sources of cyclical fluctuations across episodes. Capacity utilisation makes the largest negative contribution during the downturns of 2009, 2012–2013, and 2020, whereas the capital component plays a more prominent role in 2015–2016, 2022–2023, and the pronounced downturn in 2024Q1; the two components contribute more evenly during the shallower 2018 slowdown. Third, using firm-level quarterly data, we estimate reduced-form Phillips-curve regressions and show that the microdata-based output gap is a statistically significant and economically meaningful predictor of consumer price inflation. Fourth, we assess the real-time reliability of the sectoral output gaps directly using an expanding-sample vintage exercise. The results reveal substantial cross-sectoral heterogeneity: the manufacturing gap is only marginally affected by end-point revisions, whereas the construction and, particularly, services gaps are subject to statistically significant real-time revisions, with the real-time root-mean-square revision for services exceeding the standard deviation of the estimated gap itself. Fifth, we assess the sensitivity of the underlying production-function estimates to the choice of proxy variable in the control-function approach, re-estimating the entire framework under a Levinsohn–Petrin/ACF specification (Appendix C) as an alternative to the baseline Olley–Pakes/ACF estimator; the resulting output-gap paths are highly consistent, with correlations above 0.97 for both the aggregate and weighted measures, indicating that the main cyclical findings do not hinge on this particular identifying assumption.

The remainder of the paper is organised as follows. Section 2 presents the full production-function methodology, including the estimation of sector-specific production functions, the seasonal-adjustment and trend-extraction procedure, the derivation of the quarterly output gap and its decomposition, and the business-cycle-dating approach. Section 3 describes the quarterly data sources and sample. Section 4 presents the quarterly output-gap estimates, their decomposition, the business-cycle analysis and reports the Phillips-curve application. Section 5 concludes, discussing the relevance of the proposed measure, its limitations, and directions for future research.

2 Methodology

2.1 Production function with capacity utilisation

We assume that the gross value added Y_{it} of firm i in period t is generated by a sector-specific Cobb–Douglas production function

$$Y_{it} = TFP^{A}_{it}\, U_{it}^{\beta_{u,d}}\, K_{it}^{\beta_{k,d}}\, L_{it}^{\beta_{l,d}}\, e^{\varepsilon_{it}}, \qquad i \in S_d, \tag{1}$$

where K_{it} denotes real capital, L_{it} is the number of full-time-equivalent employees, and $U_{it} \in [0, 100]$ is the firm's self-reported capacity utilisation from the BTS ("At what capacity is your company currently operating, as a percentage of full capacity?"). TFP^{A}_{it} denotes utilisation-adjusted total factor productivity (purified technology), while $e^{\varepsilon_{it}}$ captures an output shock. The parameters $\beta_{u,d}$, $\beta_{k,d}$, and $\beta_{l,d}$ are sector-specific (d = NACE Rev. 2 division) output elasticities of capacity utilisation, capital, and labour, respectively.

Total factor productivity, $TFP_{it} = Y_{it} / \left(K_{it}^{\beta_{k,d}} L_{it}^{\beta_{l,d}}\right)$, can be decomposed into utilisation-adjusted productivity and the contribution of capacity utilisation according to $TFP_{it} = TFP^{A}_{it} \cdot U_{it}^{\beta_{u,d}}$. We assume that TFP^{A}_{it} follows a Markov process

$$TFP^{A}_{it} = e^{\xi_{it}} g\left(TFP^{A}_{i,t-1}, \lambda_{r,d}, \lambda_{t,d}\right), \qquad i \in S_d, \tag{2}$$

where $\lambda_{r,d}$ and $\lambda_{t,d}$ denote region and time (quarter) dummies, respectively, and $e^{\xi_{it}}$ represents an idiosyncratic productivity innovation.

Direct ordinary-least-squares estimation of Equation (1) is biased because unobserved utilisation-adjusted productivity TFP^{A}_{it} is correlated with the firm's endogenously chosen inputs: more productive firms tend to employ more capital and labour and operate at higher capacity utilisation. We address this endogeneity using the Olley and Pakes (1996) or Levinsohn and Petrin (2003) control-function method with the Ackerberg–Caves–Frazer (ACF) correction (Ackerberg et al., 2015). The baseline model is the Olley–Pakes specification, with firm-level investment outlays serving as the proxy for unobserved productivity.

Estimation proceeds in two stages separately for each sector S_d. In the first stage, unobserved productivity is approximated by a low-order polynomial in capital, labour, capacity utilisation, and the investment proxy. Investment decisions provide information about firm productivity that is observed by the firm but not by the econometrician. This stage identifies the sum of the input coefficients but not $\beta_{u,d}$, $\beta_{k,d}$, and $\beta_{l,d}$ individually. Following the ACF timing convention, labour and capacity utilisation are assumed to be chosen by the firm *before* the realisation of the current-period productivity shock $e^{\xi_{it}}$, while capital is a state variable fixed at the start of the period.

In the second stage, the coefficients $\beta_{u,d}$, $\beta_{k,d}$, and $\beta_{l,d}$ are recovered by the generalised

method of moments (GMM), exploiting the Markov property of TFP_{it}^{A} in Equation (2). Current capital, lagged labour, lagged capacity utilisation, and the lagged investment proxy are used as instruments, as they are assumed to be uncorrelated with the current-period productivity innovation $e^{\xi_{it}}$ (see Rovigatti and Mollisi, 2018, and Appendix B for the full moment condition). Standard errors of $\hat{\beta}_{u,d}$, $\hat{\beta}_{k,d}$, and $\hat{\beta}_{l,d}$ are obtained by a block bootstrap over firms, resampling entire firm histories to preserve the panel's serial-correlation structure.

This two-stage procedure is applied separately to each two-digit NACE division within manufacturing, construction, and services, allowing the input elasticities to vary freely across divisions. As a robustness check, Appendix C compares the resulting quarterly output gap with that obtained under the alternative Levinsohn–Petrin/ACF (LP-ACF) specification.

2.2 Seasonal adjustment and trend extraction of firm-level inputs

Firm-level quarterly flow variables (gross value added Y_{it}, labour L_{it}) and the capacity-utilisation series U_{it} exhibit pronounced and heterogeneous seasonal patterns across NACE divisions (e.g. construction activity is strongly seasonal, while several service divisions exhibit little or no seasonality). Before extracting potential (trend) input levels, we therefore seasonally adjust each firm-level series. We implement a two-step procedure:

(i) Missing quarterly observations for a given firm are first imputed by linear interpolation between the firm's own observed values; no extrapolation is performed outside the observed range. A firm therefore contributes to the sample only over the span of quarters in which it is observed in the survey.

(ii) Each resulting firm-level quarterly series is then seasonally adjusted using an automatically selected ETS (Error–Trend–Seasonal) exponential smoothing model, following the state-space framework of Hyndman et al. (2002). The ETS framework allows for alternative specifications of the error, trend, and seasonal components, with the appropriate model selected separately for each firm-level series using statistical model-selection criteria. For models with multiplicative seasonality, observations are seasonally adjusted by dividing them by the estimated seasonal component; for models with additive seasonality, the estimated seasonal component is subtracted from the observations. If no seasonal component is selected, the series is left unchanged. This procedure yields the seasonally adjusted firm-level series L_{it}^{SA}, U_{it}^{SA}, and Y_{it}^{SA}.

The seasonally adjusted firm-level series are then filtered to extract the potential (long-run trend) levels of labour, capital[1], and capacity utilisation, denoted $L_{it}^{SA,POT}$, $K_{it}^{SA,POT}$,

[1] Capital is a stock variable, measured at the beginning of the quarter, and is therefore not seasonally

and $U_{it}^{SA,POT}$. We apply the boosted Hodrick–Prescott (HP) filter (Hodrick and Prescott, 1997; Phillips and Shi, 2021) separately to each firm's individual time series, using the standard smoothing parameter for quarterly data ($\lambda = 1600$). The boosting procedure of Phillips and Shi (2021) iterates the conventional HP filter, with the number of iterations determined by a data-driven stopping rule (the Bayesian information criterion, BIC, in our implementation). This procedure is designed to reduce the residual low-frequency variation that may remain in the cyclical component after a single application of the HP filter and to improve trend estimation relative to the conventional one-shot HP filter used in the annual framework of Błażej et al. (2025). These properties are particularly relevant in the present quarterly setting, where the reliability of trend and output-gap estimates near the end of the sample is an important concern.

2.3 Potential output and the quarterly output gap

Substituting the seasonally adjusted potential inputs into the estimated production function and taking logs yields the following firm-level potential outputs at quarter t:

$$y_{it}^{POT,UKL} = \hat{\beta}_{u,d}\, u_{it}^{SA,POT} + \hat{\beta}_{k,d}\, k_{it}^{SA,POT} + \hat{\beta}_{l,d}\, l_{it}^{SA,POT} + \omega_{it}, \tag{3}$$

$$y_{it}^{POT,U} = \hat{\beta}_{u,d}\, u_{it}^{SA,POT} + \hat{\beta}_{k,d}\, k_{it}^{SA} + \hat{\beta}_{l,d}\, l_{it}^{SA} + \omega_{it}, \tag{4}$$

$$y_{it}^{POT,K} = \hat{\beta}_{u,d}\, u_{it}^{SA} + \hat{\beta}_{k,d}\, k_{it}^{SA,POT} + \hat{\beta}_{l,d}\, l_{it}^{SA} + \omega_{it}, \tag{5}$$

$$y_{it}^{POT,L} = \hat{\beta}_{u,d}\, u_{it}^{SA} + \hat{\beta}_{k,d}\, k_{it}^{SA} + \hat{\beta}_{l,d}\, l_{it}^{SA,POT} + \omega_{it}, \tag{6}$$

where lower-case letters denote logarithms and $\omega_{it} = \log TFP_{it}^{A} + \varepsilon_{it}$ is the utilisation-adjusted productivity term implied by Equation (1), recovered as the firm-level residual $\omega_{it} = y_{it}^{SA} - \hat{\beta}_{u,d} u_{it}^{SA} - \hat{\beta}_{k,d} k_{it}^{SA} - \hat{\beta}_{l,d} l_{it}^{SA}$. Holding ω_{it} fixed when actual inputs are replaced by their potential counterparts is what makes the decomposition exact: the productivity term cancels, so the firm-level gaps depend only on the deviations of inputs from their trends and on the estimated elasticities. The corresponding firm-level output gaps are $x_{it}^{UKL} = y_{it}^{SA} - y_{it}^{POT,UKL}$, $x_{it}^{U} = y_{it}^{SA} - y_{it}^{POT,U}$, $x_{it}^{K} = y_{it}^{SA} - y_{it}^{POT,K}$, $x_{it}^{L} = y_{it}^{SA} - y_{it}^{POT,L}$, and the following decomposition holds:

$$x_{it}^{UKL} = \hat{\beta}_{u,d}\big(u_{it}^{SA} - u_{it}^{SA,POT}\big) + \hat{\beta}_{k,d}\big(k_{it}^{SA} - k_{it}^{SA,POT}\big) + \hat{\beta}_{l,d}\big(l_{it}^{SA} - l_{it}^{SA,POT}\big), \tag{7}$$

$$x_{it}^{U} = \hat{\beta}_{u,d}\big(u_{it}^{SA} - u_{it}^{SA,POT}\big), \tag{8}$$

$$x_{it}^{K} = \hat{\beta}_{k,d}\big(k_{it}^{SA} - k_{it}^{SA,POT}\big), \tag{9}$$

$$x_{it}^{L} = \hat{\beta}_{l,d}\big(l_{it}^{SA} - l_{it}^{SA,POT}\big), \tag{10}$$

which decomposes the firm-level output gap into capacity-utilisation-, capital-, and labour-based components.

adjusted.

To construct output-gap measures for the whole economy, individual sectors, and regions, we consider each level of aggregation $s \in \{S, d, r\}$ and construct two microdata-based measures. The aggregated measure $x_{t,s}^{A,j}$ is obtained by first summing firm-level actual and potential outputs, $Y_{it} = e^{y_{it}^{SA}}$ and $Y_{it}^{POT,j} = e^{y_{it}^{POT,j}}$, $j \in \{UKL, U, K, L\}$, across firms belonging to the entire economy (S), a given sector (d), or region (r), and then taking the log difference between the resulting aggregated actual and potential output. The weighted measure $x_{t,s}^{W,j}$ is instead constructed as the weighted sum of the corresponding firm-level output gaps in (7),

$$x_{t,s}^{A,j} = \log Y_{st}^{SA} - \log Y_{st}^{POT,j}, \tag{11}$$

$$x_{t,s}^{W,j} = \sum_{i \in s} w_{t,i}\, x_{it}^{j}, \qquad s \in \{S, d, r\},\ j \in \{UKL, U, K, L\}, \tag{12}$$

where the firm weights $w_{t,i}$ are non-negative and sum to one within each aggregate s; in the baseline implementation every firm in s receives the gross-value-added-based. The two measures therefore answer different questions: $x_{t,s}^{A,j}$ describes the cyclical position of *output* in aggregate s, whereas $x_{t,s}^{W,j}$ describes the cyclical position of the *typical firm* in s. The economy-wide quarterly output gaps $x_{t,S}^{A,UKL}, x_{t,S}^{W,UKL}$ are decomposed into the three input-based components and interaction (residual) effects,

$$x_{t,S}^{A,UKL} = x_{t,S}^{A,U} + x_{t,S}^{A,K} + x_{t,S}^{A,L} + res_t, \tag{13}$$

$$x_{t,S}^{W,UKL} = x_{t,S}^{W,U} + x_{t,S}^{W,K} + x_{t,S}^{W,L} \tag{14}$$

and analogously into weighted sectoral and regional contributions,

$$x_{t,S}^{Type,UKL} = \sum_{d} w_{t,d}\, x_{t,d}^{Type,UKL} + res_{t,\text{sectoral}}, \tag{15}$$

$$x_{t,S}^{Type,UKL} = \sum_{r} w_{t,r}\, x_{t,r}^{Type,UKL} + res_{t,\text{regional}}, \tag{16}$$

with $Type = A$ or $Type = W$ and gross-value-added-based weights $w_{t,d}$ and $w_{t,r}$.

2.4 Business-cycle dating

The quarterly microdata-based output gap provides the basis for characterising the cyclical position of the Polish economy and identifying successive phases of the business cycle. We use two complementary representations of these dynamics.

We first apply a band-pass-type Christiano–Fitzgerald filter (Christiano and Fitzgerald, 2003) to the raw quarterly microdata-based output gap to remove residual high-frequency fluctuations while retaining variation at business-cycle frequencies (typically

1.5–8 years), yielding the cyclically adjusted series $\tilde{x}_{t,S}^{A,UKL}$. This series is used to identify local peaks and troughs that delimit expansion and contraction phases according to the classical Burns–Mitchell turning-point convention (Burns and Mitchell, 1946), subject to a minimum phase length of two quarters. We also construct cyclically adjusted sectoral output-gap paths for manufacturing, construction, and services to compare their cyclical dynamics.

As a complementary representation, we construct a business-cycle clock following the phase-diagram approach used by statistical agencies (e.g. Eurostat's business-cycle clock, published as part of the Joint Harmonised EU Programme of Business and Consumer Surveys, European Commission, 2023). We plot $\tilde{x}_{t,S}^{A,UKL}$ against its own quarter-on-quarter change, $\Delta\tilde{x}_{t,S}^{A,UKL}$, in a two-dimensional phase plane divided into four quadrants – *expansion* ($\tilde{x} > 0$, $\Delta\tilde{x} > 0$), *slowdown* ($\tilde{x} > 0$, $\Delta\tilde{x} < 0$), *recession* ($\tilde{x} < 0$, $\Delta\tilde{x} < 0$), and *recovery* ($\tilde{x} < 0$, $\Delta\tilde{x} > 0$). The resulting business-cycle clock provides a compact representation of the economy's cyclical position and direction of change as it moves through successive phases of the cycle.

For each identified peak-to-trough contraction and trough-to-peak expansion, we calculate its duration in quarters. We also calculate peak-to-peak and trough-to-trough durations as measures of full-cycle length and report summary statistics for phase and cycle durations over the sample period.

2.5 Macroeconomic validation: the Phillips curve

To assess the macroeconomic relevance of the quarterly microdata-based output gap, we estimate a standard backward-looking, reduced-form Phillips curve:

$$\pi_t = \delta_0 + \delta_\pi \pi_{t-1} + \delta_x x_t + \varepsilon_t, \tag{17}$$

where π_t denotes the quarterly consumer price index (corresponding quarter of the previous year = 100) and x_t denotes, alternately, the aggregate or weighted microdata-based output gap (boosted and unboosted versions) or a conventional Hodrick–Prescott-filtered gap of Polish GDP. All specifications are estimated over 2008Q1–2025Q1 using GMM with macroeconomic instruments (see Section 4.5 for the full specification and results). The Phillips-curve exercise provides a macroeconomic validation of the proposed measure by examining whether the microdata-based output gap contains information about inflationary pressures consistent with the expected relationship between cyclical activity and inflation.

The quarterly output gap may also have applications beyond inflation analysis. In particular, it can serve as an input into the calculation of a cyclically adjusted (structural) fiscal balance, a standard component of EU fiscal surveillance. We identify this as a potential extension of the framework but do not implement it quantitatively in the present

paper.

3 Data

3.1 Quarterly Non-Financial Enterprises Survey and Business Tendency Survey

The quarterly analysis combines two firm-level datasets collected by Statistics Poland:

(i) the *Quarterly Non-Financial Enterprises Survey* (QNFES), which is a census of approximately 17,500 enterprises with 50 or more employees per quarter and provides information on gross value added, capital, labour, and investment;

(ii) the quarterly *Business Tendency Survey* (BTS), covering approximately 11,100 firms per quarter, which provides a direct, survey-based measure of capacity utilisation together with a range of qualitative assessments of business conditions (see Statistics Poland, 2023; European Commission, 2023).

Matching QNFES and BTS respondents by their REGON identification number yields a final unbalanced panel of approximately 5,300 firms per quarter (around 10,000 unique entities) spanning 2008Q1–2025Q1, that is 69 quarters. The sample covers manufacturing (NACE section C), construction (section F), and services, while excluding wholesale and retail trade, mining and quarrying, and energy, gas, and water supply, for which capacity-utilisation data are not collected in the BTS. After these exclusions, 62 two-digit NACE divisions remain: 23 in manufacturing, 3 in construction, and 36 in services. Because several of the service divisions retained in the sample are only partly market-determined (for example divisions 85 and 86), we refer to this group simply as services rather than as market services.

Because QNFES is a census of firms with 50 or more employees, the quarterly sample is concentrated among medium and large enterprises and does not capture the cyclical behaviour of micro and small firms directly. Nevertheless, medium and large enterprises account for the majority of sectoral gross value added in Poland, so the matched sample captures a substantial share of economic activity in the sectors covered by the analysis. The resulting output gap should therefore be interpreted primarily as a measure of cyclical conditions among medium and large enterprises, while remaining informative about broader sectoral dynamics.

3.2 Sample construction

Firm-level gross value added and capital are expressed in constant prices using two-digit NACE sector-specific deflators (see Appendix A for the exact definitions of the under-

lying accounting variables), while labour is measured in full-time-equivalent employees. Missing observations within each firm's observed sample period are imputed by linear interpolation, with no extrapolation beyond the firm's first and last observed quarters. Capacity utilisation is measured as a percentage of full capacity, $U_{it} \in [0, 100]$. Table 1 summarises the two datasets and the resulting matched sample.

Table 1: Data sources: Poland, quarterly sample, 2008Q1–2025Q1

Dataset	Coverage	Key variables
QNFES	∼17,500 firms/quarter, census of firms with 50+ employees	GVA, capital, labour, investment
BTS	∼11,100 firms/quarter	Capacity utilisation, business assessments
Final matched sample (QNFES∩BTS)		
∼5,300 firms/quarter; ∼10,000 unique entities; unbalanced panel, 2008Q1–2025Q1		

4 Empirical results

4.1 Production function estimates

Table 2 reports summary statistics for the estimated input elasticities $\hat{\beta}_{u,d}$, $\hat{\beta}_{k,d}$, and $\hat{\beta}_{l,d}$, as well as for the implied returns to scale $\hat{\beta}_{k,d} + \hat{\beta}_{l,d}$, across the 62 two-digit NACE divisions for which the quarterly production function (1) is estimated. These comprise 23 manufacturing divisions, 3 construction divisions, and 36 service divisions. Figure 1 presents the full set of division-level estimates.

Table 2: Summary of quarterly production-function input elasticities by sector, 2008Q1–2025Q1

		Mean (standard deviation)			Range [min, max]		
Sector	N_d	$\hat{\beta}_{l,d}$	$\hat{\beta}_{k,d}$	$\hat{\beta}_{u,d}$	$\hat{\beta}_{l,d}$	$\hat{\beta}_{k,d}$	$\hat{\beta}_{u,d}$
Manufacturing (C)	23	0.786 (0.164)	0.272 (0.119)	0.276 (0.106)	[0.402, 1.213]	[0.111, 0.558]	[0.068, 0.497]
Construction (F)	3	0.943 (0.049)	0.159 (0.043)	0.183 (0.051)	[0.904, 0.998]	[0.109, 0.189]	[0.145, 0.241]
Services	36	0.812 (0.167)	0.169 (0.136)	0.158 (0.196)	[0.464, 1.135]	[−0.069, 0.575]	[−0.373, 0.636]
All divisions (S)	62	0.809 (0.164)	0.207 (0.135)	0.203 (0.172)	[0.402, 1.213]	[−0.069, 0.575]	[−0.373, 0.636]

Notes: N_d is the number of two-digit NACE divisions in the sector. Sector-level statistics are unweighted averages across divisions. Returns to scale are defined with respect to capital and labour only, $\hat{\beta}_{k,d} + \hat{\beta}_{l,d}$, since capacity utilisation is an intensity of use of the installed capital stock rather than a separately accumulated input. They average 1.059 (manufacturing), 1.102 (construction), 0.981 (services), and 1.016 across all 62 divisions, close to constant returns to scale. Estimates are obtained by the two-stage Olley–Pakes/ACF control-function procedure of Section 2.1, applied separately to each division on quarterly firm-level data, 2008Q1–2025Q1.

Three main patterns emerge. First, manufacturing divisions exhibit the highest aver-

age capital elasticity (0.272), compared with 0.159 in construction and 0.169 in services, consistent with the relatively capital-intensive nature of manufacturing production. The highest capital elasticity within manufacturing is estimated for division 11 (manufacture of beverages, $\hat{\beta}_k = 0.558$), a capital-intensive activity, consistent with firm-level capital-elasticity estimates for Polish manufacturing reported elsewhere (Górajski and Błażej, 2020). Average returns to scale in manufacturing are close to constant returns (1.059), while construction exhibits the highest average returns to scale among the three sectors (1.102).

Second, construction divisions form the most homogeneous group in terms of the estimated elasticities, with cross-division standard deviations of approximately 0.05 or less, compared with substantially greater dispersion in manufacturing and services. Construction also exhibits the highest average labour elasticity (0.943), compared with 0.786 in manufacturing and 0.812 in services, consistent with the relatively labour-intensive nature of construction activity.

Third, service divisions display the greatest cross-division heterogeneity, particularly in the estimated capital and capacity-utilisation elasticities. Capital elasticities range from -0.069 (division 72, scientific research and development) to 0.575 (division 77, rental and leasing), while capacity-utilisation elasticities range from -0.373 (division 85, education) to 0.636 (division 91, libraries, archives, museums). The comparatively wide dispersion of the capacity-utilisation estimates suggests that the survey-based concept of capacity utilisation may capture differences in productive-capacity use less uniformly across service activities than in manufacturing or construction.

Capacity utilisation has the highest average output elasticity in manufacturing (0.276), compared with 0.183 in construction and 0.158 in services. This result is consistent with capacity utilisation playing a relatively important role in short-run output adjustment in manufacturing, where production can respond through changes in machine use, operating hours, and shift intensity. A small number of service divisions display negative capacity-utilisation or capital elasticities. These estimates may reflect both the relatively small number of participating firms in some divisions of the matched QNFES∩BTS sample and limitations in applying a survey-based capacity-utilisation concept to activities in which productive capacity is less closely related to physical capital, such as education. These division-level estimates should therefore be interpreted with caution, although their influence on the economy-wide output gap is limited by the gross-value-added weighting used in the aggregation described in Section 2.3.

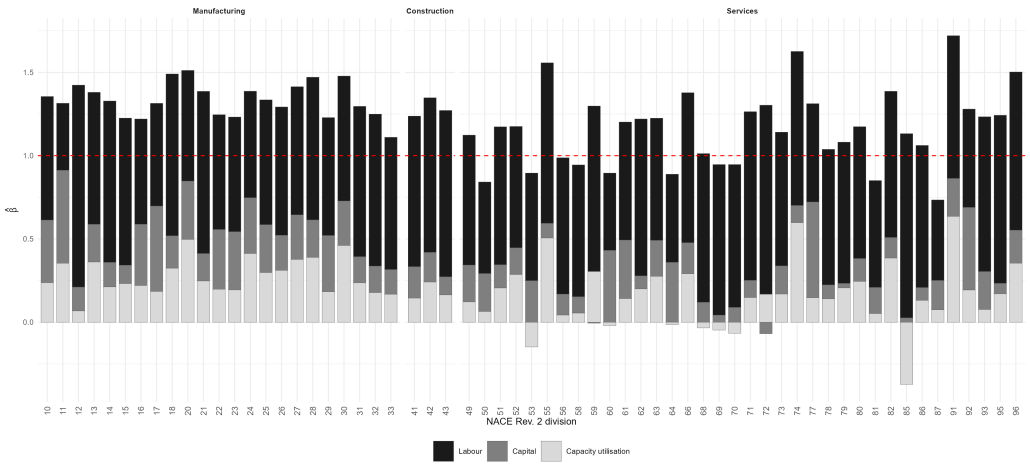


Figure 1: Quarterly production-function input elasticities by NACE two-digit division, 2008Q1–2025Q1 (black: labour, $\hat{\beta}_{l,d}$; dark grey: capital, $\hat{\beta}_{k,d}$; light grey: capacity utilisation, $\hat{\beta}_{u,d}$; dashed red line: constant returns to scale in capital and labour, $\hat{\beta}_{k,d} + \hat{\beta}_{l,d} = 1$). Bars are stacked, so the total bar height equals $\hat{\beta}_{l,d} + \hat{\beta}_{k,d} + \hat{\beta}_{u,d}$ and exceeds the reference line by the capacity-utilisation elasticity; the reference line should therefore be read against the sum of the two lower segments only.

4.2 The quarterly microdata-based output gap and its decomposition

Figure 2 plots the quarterly aggregate output gap $x_{t,S}^{A,UKL}$, defined as the log difference between aggregated actual and potential output, together with its decomposition into capacity-utilisation, capital, and labour components (Equation (13)). The weighted output gap and its decomposition are presented in the Appendix D. Although the raw quarterly series is considerably more volatile than its cyclically adjusted counterpart discussed in Section 4.4, it reveals the same six broad cyclical episodes, which are formally dated later in that section.

The first episode corresponds to the Global Financial Crisis. The output gap falls to −2.51% in 2009Q2, largely due to a decline in capacity utilisation (−1.56 percentage points), before a brief recovery and a renewed deterioration to −2.99% in 2010Q1, the lowest value of this episode. Capacity utilisation remains the main contributor (−1.69 pp), accompanied by a substantial negative labour contribution (−1.02 pp). The subsequent 2012–2013 slowdown culminates in a trough of −2.67% in 2013Q1, with capacity utilisation again accounting for the largest share of the decline (−1.49 pp) and labour making an additional negative contribution (−0.73 pp).

The recovery that follows brings the output gap to +1.55% in 2015Q1, the highest positive quarterly reading in the sample, with the increase largely attributable to capital (+1.61 pp). This expansion is followed by a relatively short capital-driven contraction, with the gap falling to −1.79% in 2015Q4 and the capital component accounting for

−1.30 pp. During the subsequent 2017–2019 expansion, fluctuations are more moderate: a shallow dip to −0.63% in 2018Q2 reflects nearly equal negative contributions from capacity utilisation (−0.33 pp) and capital (−0.29 pp).

The COVID-19 shock marks the next pronounced downturn, with the output gap reaching −2.59% in 2020Q3, primarily due to the capacity-utilisation component (−1.73 pp). This is followed by a relatively rapid recovery, culminating in a post-pandemic peak of +0.29% in 2021Q4, supported mainly by capital (+0.84 pp). The final episode begins with a capital-driven contraction associated with the energy-price shock and monetary tightening. The gap declines to −1.33% in 2022Q4, almost entirely reflecting the capital component (−1.36 pp), while the contributions of labour and capacity utilisation remain close to zero. After a temporary improvement, the gap falls sharply to −4.48% in 2024Q1, the lowest quarterly reading in the entire sample, with capital alone contributing −3.12 pp. The gap subsequently recovers rapidly, returning to slightly positive territory (+0.21%) by 2025Q1.

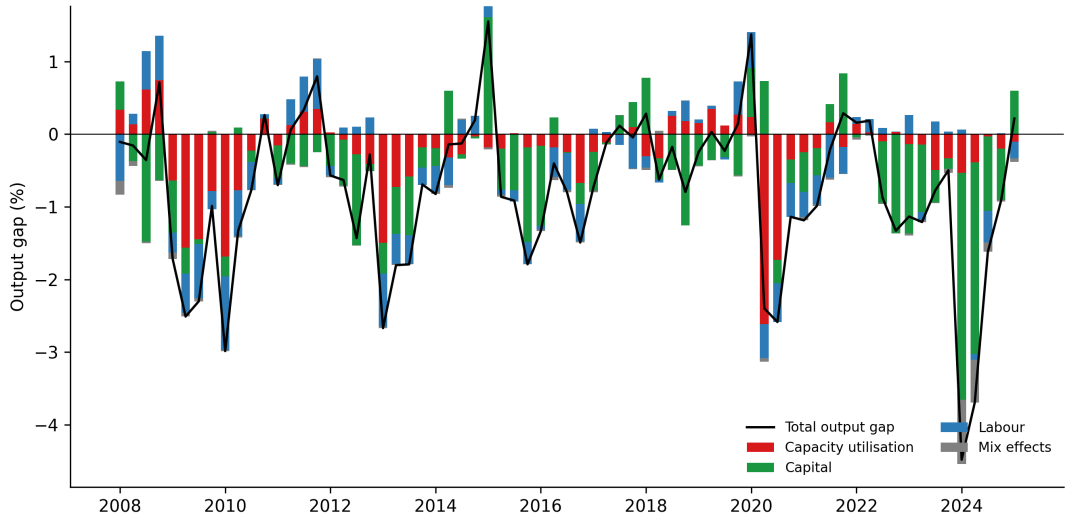


Figure 2: Aggregate microdata-based output gap $x_{t,S}^{A,UKL}$ in Poland, 2008Q1–2025Q1, decomposed into capital-, labour-, and capacity-utilisation-based components.

Figure 3 compares the raw quarterly aggregated output gap for the whole economy with the corresponding gaps for manufacturing, construction, and services.[2] Construction is the most volatile of the three sectors, with a standard deviation of 1.74 percentage points over the sample, compared with 1.43 pp for manufacturing and 0.68 pp for services. It records a trough of −5.99% in 2019Q3 – the deepest sectoral reading in the sample outside the 2024 episode – and reaches a pronounced peak of +2.03% in 2025Q1. Manufacturing tracks the total gap most closely, with a correlation of 0.95, consistent with

[2]The total plotted in Figure 3 is the gross-value-added-weighted sum of the three sectoral gaps in Equation (15); it therefore differs slightly from the economy-wide aggregate of Figure 2, which also includes the sectoral interaction term $res_{t,\text{sectoral}}$.

its relatively large weight in the aggregate measure. It records the highest single-quarter sectoral gap in the sample (+2.67% in 2015Q1), coinciding with the peak in the total gap (+1.52%), as well as a marked COVID-19 trough (−2.73% in 2020Q3), which also coincides with the total trough for that episode (−2.19%). The deepest reading of the whole sample, however, is manufacturing's 2024Q1 trough, which is substantially lower than any other sectoral observation and drives the corresponding trough in the economy-wide series documented above. Services, by contrast, displays considerably lower volatility, with a standard deviation of 0.68 pp, and is markedly less synchronised with the other series: its correlation with the total gap is 0.59, compared with 0.95 for manufacturing and just 0.30 for construction, while its correlations with manufacturing and construction are 0.37 and 0.13, respectively. Construction's low correlation with the total gap (0.30), despite being the most volatile sector, indicates that its cycle is largely idiosyncratic and only loosely synchronised with the economy-wide cycle. Overall, the services gap exhibits more muted fluctuations and, together with construction, a markedly weaker co-movement with the aggregate and with each other than manufacturing displays with the aggregate gap.

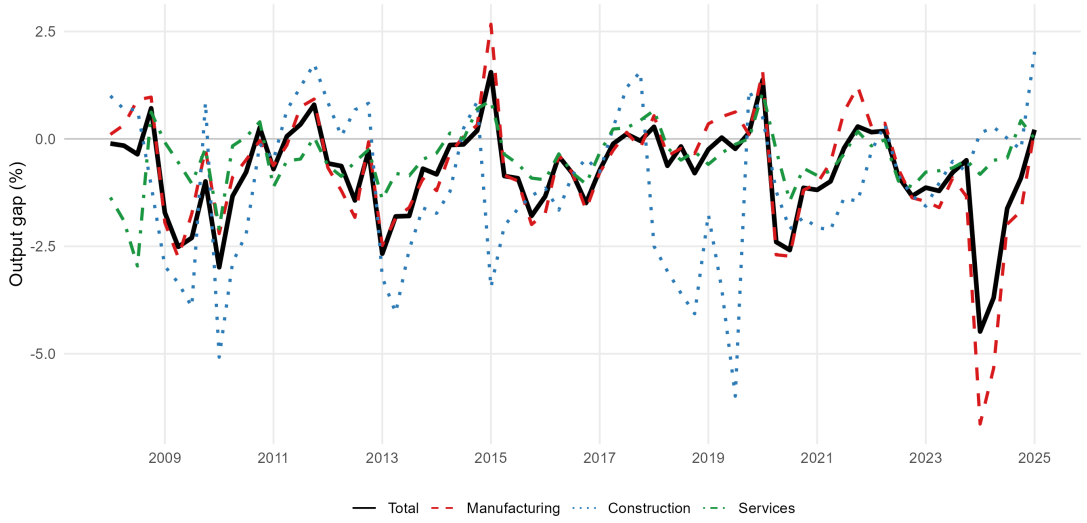


Figure 3: Aggregated microdata-based output gap by sector: total, manufacturing, construction, and services, 2008Q1–2025Q1.

4.3 Sensitivity of the output-gap estimates to sample size: a vintage exercise

We evaluate how the microdata-based output-gap estimates respond to the arrival of new data points, using nine expanding-sample vintages – 2023Q1 ($T = 61$ quarters) to 2025Q1 ($T = 69$) – for three aggregates of the Polish economy: manufacturing, construction and services. The ninth and largest vintage serves as the benchmark against which the remaining eight are evaluated, so each of the tables below reports eight rows.

For each sector $s \in \{C, F, \text{Services}\}$ let x_t^v denote the microdata-based output gap

estimated for quarter t from the information set available at vintage v. Each sector has $V = 9$ vintages, $v \in \{2023\text{Q1}, \ldots, 2025\text{Q1}\}$, every one an expanding sample beginning in 2008Q1 and ending in the quarter that names it:

$$T_{2023\text{Q1}} = 61, \;\; T_{2023\text{Q2}} = 62, \;\; \ldots, \;\; T_{2025\text{Q1}} = 69. \tag{18}$$

Consecutive vintages differ by exactly one new quarterly observation. The three sectors share the same calendar and the same vintage structure, which makes the cross-sector comparison in Section 4.3.3 exact rather than approximate: every statistic below is computed on the same $61 + 62 + \cdots + 68 = 516$ vintage–quarter cells per sector, or 3×516 in total.

The largest-sample vintage is the pseudo-final benchmark, $x_t^F \equiv x_t^{2025\text{Q1}}$, and the object of interest is the *total revision*

$$r_t^v \;=\; x_t^F - x_t^v, \qquad t = 1, \ldots, T_v, \tag{19}$$

alongside the *marginal* revision $d_t^v = x_t^{v+1} - x_t^v$, which isolates the effect of one single extra observation. Because each vintage's own sample end moves with v, revisions must be indexed by distance from the end of the estimation sample,

$$h \;=\; T_v - t, \tag{20}$$

so that $h = 0$ is the genuinely real-time estimate. This distinction does the analytical work in what follows: a statistic computed over calendar time confounds the end-point problem with data revisions, whereas the pair (calendar date, h) separates them.

For each vintage we report the mean revision MR (zero under unbiasedness), the mean absolute revision MAR, the RMSE $\left(T_v^{-1} \sum_t (r_t^v)^2\right)^{1/2}$ and the correlation ρ with the benchmark. Two evaluation windows are used throughout. The *full-overlap* window ($t \leq T_v$) mixes two things: longer samples, but also more end-of-sample observations entering the comparison. The *common* window fixes the evaluation period at 2008Q1–2023Q1 for every vintage, so that differences across vintages isolate what the extra data points do to the *historical* path of the gap; for the 2023Q1 vintage the two windows coincide by construction. It is the common window that answers the sample-size question posed here; the full-overlap window is reported for comparability with the revision literature.

4.3.1 Manufacturing

Manufacturing exhibits the greatest real-time stability among the three sectors, and Table 3 illustrates this pattern. Over the common window, the historical path of the gap generally converges towards the benchmark: MAR declines from 0.139 pp in the 2023Q1 vintage to 0.079 pp in 2024Q4, while the RMSE falls from 0.238 to 0.098 pp – a reduc-

tion of about 60% – and ρ rises from 0.987 to 0.998. The convergence is not monotonic, however. The 2024Q1 and 2024Q3 vintages temporarily interrupt this pattern, with common-window RMSEs of 0.250 and 0.232 pp, compared with 0.218 and 0.093 pp in the respective preceding vintages. Thus, although the historical estimates become increasingly stable as additional observations accumulate, individual data releases can still generate temporary revisions to the estimated path.

A different pattern emerges when the full-overlap window is considered. The RMSE increases from 0.238 to 0.293 pp and reaches a maximum of 0.375 pp in 2024Q2, while MAR remains within the relatively narrow range of 0.137–0.155 pp. Since the difference between the common and full-overlap windows arises from the additional quarters included after 2023Q1, the deterioration in the full-overlap statistics is concentrated in the newly added end-of-sample observations rather than in substantial revisions to the earlier historical path. Figure 4 illustrates this pattern: the min–max envelope remains very narrow over most of the sample and widens primarily near the end of each vintage – exceeding 1 pp around 2023Q4–2024Q1 – before narrowing again as two or three additional observations become available.

Overall, these revisions remain small relative to a gap with a standard deviation of 1.50 pp. The correlation with the benchmark never falls below 0.9757, and the sign of the gap agrees with the benchmark in 96.9% of all vintage–quarter cells, and in 88% of the eight end-point estimates. Importantly, the real-time estimate ($h = 0$) shows *no* statistically significant systematic bias: the mean end-point revision is -0.37 pp with $t = -1.15$ ($p = 0.2876$). Manufacturing therefore emerges as the sector with the greatest real-time reliability: although its most recent estimates are still subject to revision, these revisions are relatively small and show no evidence of a systematic tendency to revise the gap in one direction.

Table 3: Revisions of the aggregated microdata-based output gap, NACE Section C.

		Full overlap $t \leq T_v$			Common window $t \leq$ 2023Q1			
Vintage	T_v	MR	MAR	RMSE	MR	MAR	RMSE	ρ
2023Q1	61	0.073	0.139	0.238	0.073	0.139	0.238	0.9872
2023Q2	62	0.025	0.142	0.246	0.042	0.128	0.212	0.9883
2023Q3	63	0.039	0.142	0.243	0.063	0.123	0.209	0.9893
2023Q4	64	0.001	0.155	0.295	0.043	0.122	0.218	0.9876
2024Q1	65	0.032	0.139	0.271	0.043	0.119	0.250	0.9834
2024Q2	66	0.045	0.148	0.375	0.007	0.078	0.093	0.9976
2024Q3	67	0.027	0.152	0.259	0.047	0.131	0.232	0.9859
2024Q4	68	0.030	0.137	0.293	0.011	0.079	0.098	0.9976

Note: each vintage is evaluated against the 2025Q1 benchmark, $r_t^v = x_t^F - x_t^v$, so a positive MR means the vintage understated the gap. MR, MAR and RMSE in percentage points; ρ is the correlation with the benchmark on the common window. For the 2023Q1 vintage the two windows coincide by construction.

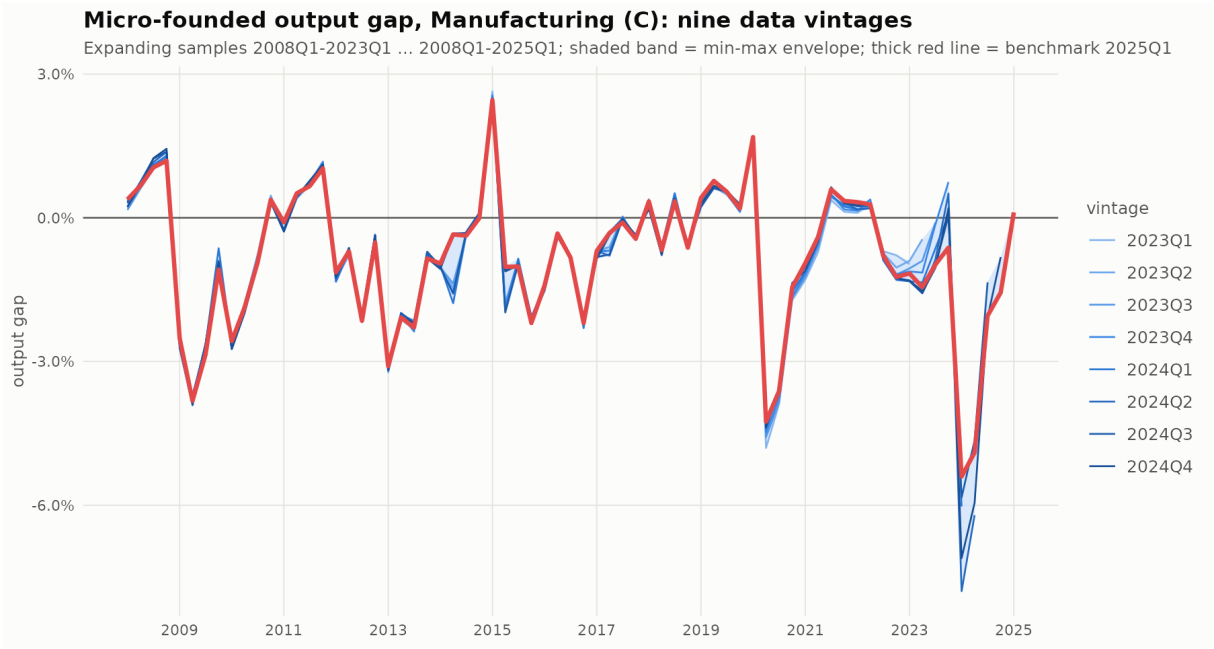


Figure 4: NACE Section C: nine expanding-sample vintages of the aggregated microdata-based output gap. Shaded band: min–max envelope across vintages; thick red line: 2025Q1 benchmark.

4.3.2 Construction

Construction exhibits the most volatile output gap among the three sectors (s.d. 1.98 pp) and the largest absolute revisions (Table 4). Two features distinguish its revision pattern from that of manufacturing. First, the revisions are consistently negative: MR is below zero in every vintage and over both evaluation windows, ranging from -0.142 to -0.042 pp on the full-overlap window. Thus, each earlier vintage places the construction gap above

the level subsequently indicated by the benchmark, implying a systematic downward revision as additional data become available.

Second, additional observations do not lead to a sustained convergence of the historical path towards the benchmark. Over the common window, the revisions initially decline as expected: MAR falls from 0.184 to 0.142 pp and the RMSE from 0.327 to 0.194 pp between the 2023Q1 and 2023Q4 vintages, while ρ rises to 0.996. This improvement is not sustained in the 2024Q2–2024Q4 vintages, however, when the RMSE increases again to 0.33–0.38 pp and ρ falls to approximately 0.986. The arrival of additional observations in 2024 therefore generates further revisions to the pre-2023 historical path rather than a continued convergence towards the benchmark. Figure 5 illustrates this pattern: the min–max envelope remains relatively wide over 2021–2024 instead of progressively narrowing towards the benchmark. The full-overlap statistics provide a similar picture. The RMSE ends at 0.334 pp, compared with 0.327 pp in the initial vintage, and reaches a maximum of 0.436 pp in 2023Q4.

The real-time estimates exhibit correspondingly substantial revisions. The real-time RMSE is 1.47 pp – 0.74 times the standard deviation of the gap itself – and the mean end-point revision is -1.21 pp with $t = -3.84$ ($p = 0.0064$), indicating a statistically significant systematic difference between the real-time and benchmark estimates. Consistent with the negative mean revision, the real-time estimates tend to indicate a stronger cyclical position in construction than is subsequently shown by the benchmark. The real-time estimate places the gap above zero in seven of the eight quarters examined, whereas the benchmark places it below zero in six of eight; consequently, the sign of the real-time and benchmark estimates coincides in only 38% of cases.

Table 4: Revisions of the aggregated microdata-based output gap, NACE Section F.

		Full overlap $t \leq T_v$			Common window $t \leq$ 2023Q1			
Vintage	T_v	MR	MAR	RMSE	MR	MAR	RMSE	ρ
2023Q1	61	-0.089	0.184	0.327	-0.089	0.184	0.327	0.9876
2023Q2	62	-0.114	0.212	0.350	-0.088	0.188	0.279	0.9912
2023Q3	63	-0.083	0.202	0.302	-0.047	0.170	0.220	0.9942
2023Q4	64	-0.142	0.209	0.436	-0.071	0.142	0.194	0.9963
2024Q1	65	-0.071	0.173	0.306	-0.025	0.134	0.199	0.9952
2024Q2	66	-0.085	0.229	0.400	-0.078	0.219	0.382	0.9837
2024Q3	67	-0.044	0.219	0.362	-0.049	0.203	0.353	0.9859
2024Q4	68	-0.042	0.215	0.334	-0.017	0.209	0.327	0.9873

Note: see Table 3. A negative MR means the vintage overstated the gap relative to the 2025Q1 benchmark.

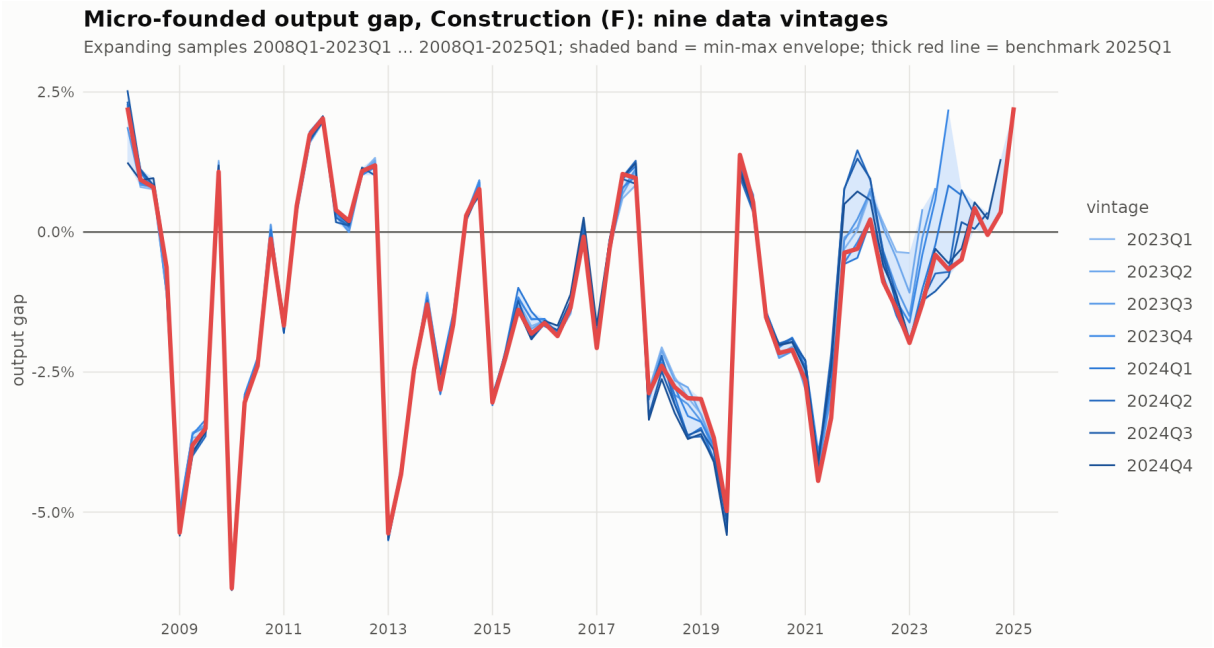


Figure 5: NACE Section F: nine expanding-sample vintages of the aggregated microdata-based output gap. Shaded band: min–max envelope across vintages; thick red line: 2025Q1 benchmark.

4.3.3 Services

Services exhibit the smallest cyclical amplitude among the three sectors (s.d. 0.75 pp), making revisions of a given absolute magnitude relatively large compared with the variation in the estimated gap. Although the absolute end-point RMSE, 1.24 pp, is smaller than that for construction, relative to the standard deviation of the gap itself it amounts to 1.65. Thus, *the magnitude of the real-time revision exceeds the cyclical variation in the estimated gap*. Sign concordance at the end point is only 25%, further indicating substantial uncertainty in the real-time assessment of the sector's cyclical position. This pattern is consistent with a broader literature documenting that services-sector price and quantity dynamics are comparatively difficult to capture in reduced-form macroeconomic models, including Phillips-curve specifications estimated on services price inflation (Tallman and Zaman, 2017).

Table 5 nevertheless reveals a pattern that differs from those observed for manufacturing and construction: services are the only sector in which additional observations improve the estimates over *both* evaluation windows, and the improvement is nearly monotonic. On the common window, MAR falls from 0.188 to 0.083 pp and the RMSE from 0.295 to 0.117 pp, while ρ rises from 0.928 to 0.991. On the full-overlap window, the RMSE also more than halves, from 0.295 to 0.137 pp, with only a temporary reversal in 2024Q3. What distinguishes services is therefore not a lack of convergence as additional observations become available, but the relatively low initial real-time stability of the estimated historical path. With $\rho = 0.928$, the 2023Q1 vintage reproduces the benchmark path less

closely than the corresponding manufacturing or construction vintages (both ≈ 0.988). Services consequently show the largest improvement as the estimation sample is extended, and by the 2024Q2–2024Q4 vintages their common-window accuracy becomes comparable with that of manufacturing.

Table 5: Revisions of the aggregated microdata-based output gap, services.

		Full overlap $t \leq T_v$			Common window $t \leq$ 2023Q1			
Vintage	T_v	MR	MAR	RMSE	MR	MAR	RMSE	ρ
2023Q1	61	-0.010	0.188	0.295	-0.010	0.188	0.295	0.9275
2023Q2	62	-0.039	0.176	0.299	-0.013	0.152	0.218	0.9609
2023Q3	63	-0.056	0.164	0.292	-0.016	0.127	0.180	0.9733
2023Q4	64	-0.037	0.167	0.259	-0.003	0.140	0.201	0.9699
2024Q1	65	-0.069	0.143	0.195	-0.043	0.122	0.160	0.9801
2024Q2	66	-0.052	0.103	0.148	-0.032	0.087	0.111	0.9906
2024Q3	67	-0.047	0.102	0.212	-0.012	0.073	0.106	0.9915
2024Q4	68	-0.057	0.092	0.137	-0.046	0.083	0.117	0.9905

Note: see Table 3. A negative MR means the vintage overstated the gap relative to the 2025Q1 benchmark.

The aggregate revision statistics, however, conceal substantial instability at the end of the sample, as illustrated in Figure 6. The min–max envelope remains narrow before 2021 but widens to approximately 1.5 pp over 2022Q4–2024Q2, which is substantial relative to a services gap that rarely moves outside the interval $[-1.5, 1.0]$ pp. Importantly, *all eight* end-point revisions are negative, and the mean absolute revision and the absolute value of the mean revision therefore coincide, at $|-1.16|$ pp. This indicates that the real-time uncertainty is not merely associated with dispersion around the benchmark but has a pronounced systematic component. With $t = -7.08$ ($p = 0.0002$), the mean end-point revision is statistically significant, providing strong evidence of a systematic downward revision in the real-time services gap as additional observations become available.

The direction of these revisions also materially affects the assessment of the sector's cyclical position. Seven of the eight real-time estimates place the services gap above zero, whereas seven of the eight corresponding benchmark estimates place it below zero. Thus, for most of the quarters examined, the real-time and benchmark estimates imply opposite signs of the services output gap. In particular, a real-time assessment based on the available estimates in 2023–2024 would generally have indicated a positive services gap, whereas the corresponding benchmark estimates based on the extended sample are predominantly negative.

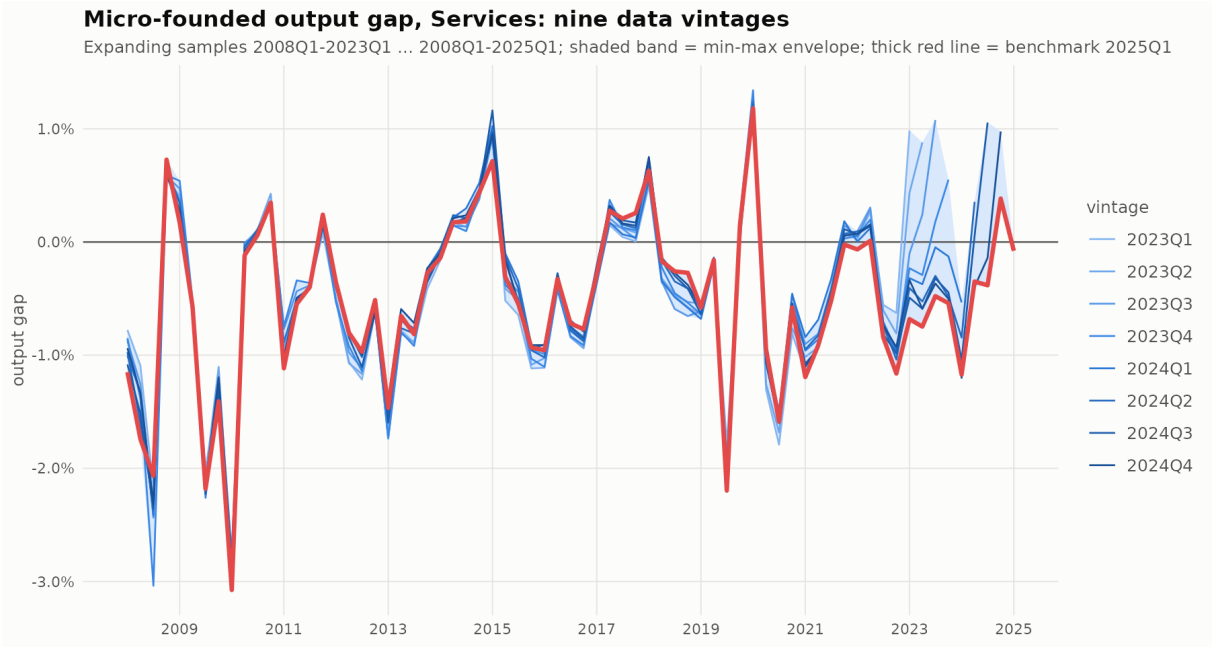


Figure 6: Services: nine expanding-sample vintages of the aggregated microdata-based output gap. Shaded band: min–max envelope across vintages; thick red line: 2025Q1 benchmark.

Cross-sector comparison. Taken together, the three sectors display markedly different real-time behaviour despite being evaluated using an identical vintage design. Manufacturing is the only sector with no statistically significant end-point bias ($p = 0.2876$) and exhibits the highest end-point sign concordance (88%). Its real-time RMSE, 0.92 pp, is the smallest of the three in absolute terms and amounts to 0.62 times the standard deviation of the gap itself; away from the end of the sample its revisions are an order of magnitude smaller still (0.21 pp, that is, 4.5 times smaller than at the end point). Construction and services, by contrast, both exhibit statistically significant systematic end-point revisions (construction: $p = 0.0064$; services: $p = 0.0002$), with revisions consistently occurring in the same direction. Their real-time RMSEs are also large relative to the underlying cyclical variation, amounting to 0.74 and 1.65 times the respective sectoral standard deviations. For services, this ratio exceeds one, implying that the typical magnitude of the real-time error is greater than the cyclical variation in the estimated gap itself. Sign concordance at the real-time horizon falls to 38% for construction and 25% for services, indicating substantial uncertainty even about the sign of the contemporaneous output gap. Overall, the central finding of the vintage exercise is the pronounced cross-sectoral heterogeneity in real-time reliability: the performance of the quarterly microdata-based output gap cannot be assumed to be uniform across sectors and should therefore be assessed separately for each sector.

4.4 Business-cycle dating and sectoral cycles

Figure 7 plots the adjusted (Christiano–Fitzgerald filtered) aggregate output gap together with the corresponding gaps for manufacturing, construction, and services. By isolating fluctuations at business-cycle frequencies, the filtered aggregate series reveals six complete peak-to-trough contractions, delimited by six peaks (2008Q1, 2011Q3, 2014Q4, 2017Q3, 2019Q3 and 2022Q1) and six troughs (2009Q4, 2013Q2, 2016Q2, 2018Q3, 2020Q4 and 2024Q1); see Table 6 below. The final trough is followed by an incipient recovery that had not yet reached a new peak by the end of the sample. Sectoral differences are pronounced. Construction displays the greatest cyclical volatility, with a standard deviation of 1.17 percentage points over the sample, compared with 0.92 pp for manufacturing and 0.38 pp for services. Services, in turn, exhibit the smoothest and most muted cyclical profile, consistent with their lower exposure to capital-intensive and tradable production. The final contraction is both deeper and more persistent than the preceding episodes and is particularly pronounced in manufacturing: the manufacturing gap falls to −2.26% in 2024Q1, well below its two preceding cyclical troughs. This pattern is consistent with the sharp, capital-driven decline in the aggregate raw series documented in Section 4.2.

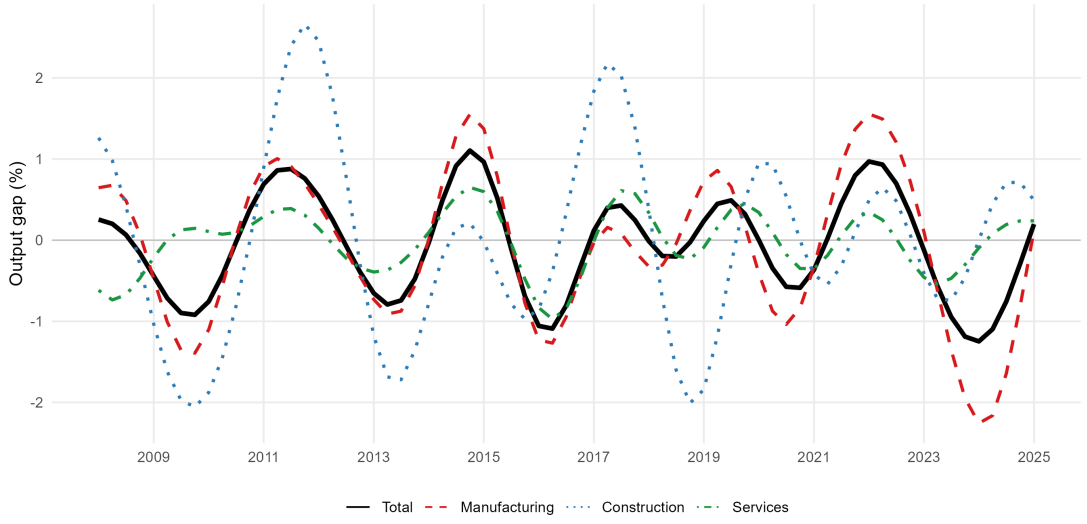


Figure 7: Cyclically adjusted aggregate microdata-based output gap by sector: total, manufacturing, construction, and services.

The sequence of cyclical phases is further illustrated by the business-cycle clock in Figure 8, which plots the cyclically adjusted aggregate output gap against its quarterly change. The Polish economy follows a broadly clockwise trajectory through the four canonical phases of the cycle – expansion, slowdown, recession, and recovery. The amplitude of these movements varies considerably over time. The loops widen around 2009, 2015–2016, and most notably in 2023–2024, when the economy experiences the largest-amplitude cycle in the sample, associated with the manufacturing-led downturn described

above. By contrast, the loops narrow during 2017–2019, indicating smaller and more moderate cyclical fluctuations.

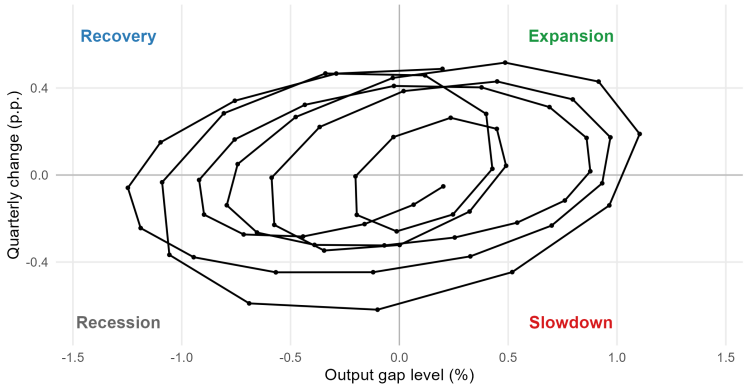


Figure 8: Business cycle clock: adjusted aggregate microdata-based output gap (horizontal axis) against its quarterly change (vertical axis), 2008Q1–2025Q1.

The timing and duration of these cyclical phases are summarised in Table 6, which reports the peak and trough dates and the duration, in quarters, of each identified contraction (peak-to-trough) and expansion (trough-to-peak), together with peak-to-peak and trough-to-trough measures of full-cycle length. All dates are determined directly from the turning points of the cyclically adjusted aggregate output gap. On average, contractions last 6.2 quarters and completed expansions 5.4 quarters, implying an average full-cycle length of approximately 11.2–11.4 quarters, or about 2.8–2.9 years, over 2008–2025. These averages, however, conceal considerable variation across episodes. The final contraction, from the 2022Q1 peak to the 2024Q1 trough, is the longest in the sample, lasting eight quarters, and corresponds to the deepest downturn observed in the raw series (Section 4.2). The subsequent recovery brings the cyclically adjusted gap to +0.20% by 2025Q1 (against +0.21% for the unfiltered series), but no new cyclical peak is identified before the end of the sample. At the other end of the spectrum, the 2017–2019 period is characterised by shorter and shallower fluctuations, including a mild contraction between 2017Q3 and 2018Q3. Such movements might fall below the amplitude thresholds adopted by some institutional business-cycle dating committees, but they satisfy the minimum two-quarter phase-length criterion applied here.

Table 6: Duration of business-cycle expansions and contractions in Poland, 2008Q1–2025Q1 (quarters), aggregate output gap

		Contraction / expansion		Full cycle	
Peak	Trough	Peak-to-trough	Trough-to-peak	Peak-to-peak	Trough-to-trough
2008Q1	2009Q4	7	7	–	–
2011Q3	2013Q2	7	6	14	14
2014Q4	2016Q2	6	5	13	12
2017Q3	2018Q3	4	4	11	9
2019Q3	2020Q4	5	5	8	9
2022Q1	2024Q1	8	*(ongoing)*	10	13
Mean duration		6.2	5.4	11.2	11.4

Notes: Peaks and troughs are local extrema of the adjusted aggregate output gap $\tilde{x}_{t,S}^{A,UKL}$ (Christiano–Fitzgerald filter), subject to a minimum phase length of two quarters. The 2008Q1 peak is a boundary observation, so the preceding expansion cannot be dated. The expansion beginning at the 2024Q1 trough had not reached a confirmed peak by 2025Q1, the end of the estimation sample, and is therefore excluded from the calculation of the mean trough-to-peak duration. The preceding peak-to-trough contraction, lasting eight quarters, is complete and is included in the corresponding mean-duration calculation.

4.5 Macroeconomic validation: the Phillips curve estimates

We estimate the reduced-form Phillips curve in Equation (17) for quarterly CPI inflation over 2008Q1–2025Q1 using the generalised method of moments (GMM) with heteroskedasticity-and-autocorrelation-consistent (HAC) weighting. To address the potential endogeneity of contemporaneous macroeconomic regressors, we use a set of predetermined macroeconomic variables as instruments: the lagged change in import prices, twice-lagged CPI inflation, the lagged unemployment rate, lagged M3 money-supply growth, lagged new manufacturing orders, the lagged change in the European Commission's Economic Sentiment Indicator, lagged real GDP growth, and the lagged output gap measure. The GMM framework also allows us to assess the validity of the overidentifying restrictions using the J-test.

We estimate five specifications, differing in the output-gap measure included in the Phillips curve. These comprise the aggregate and weighted microdata-based output gaps, $x_{t,S}^{A,UKL}$ and $x_{t,S}^{W,UKL}$, each constructed using the unboosted and boosted (Phillips and Shi, 2021) HP-filter trend specifications, together with a conventional HP-filtered ($\lambda = 1600$) aggregate Polish GDP gap used as a macroeconomic benchmark. Table 7 reports the estimated coefficients and their p-values, together with R^2 and the p-value of the GMM overidentification (J) test for each specification.

Table 7: Quarterly Phillips curve estimates, Poland, 2008Q1–2025Q1: GMM estimation of $\pi_t = \delta_0 + \delta_\pi \pi_{t-1} + \delta_x x_t + \varepsilon_t$

	Aggregate microdata-based gap		Weighted microdata-based gap		Macro benchmark
	HP	Boosted HP	HP	Boosted HP	GDP, HP
$\hat{\delta}_0$	29.60**	19.52*	33.75**	22.43	36.02***
	(0.022)	(0.093)	(0.014)	(0.152)	(0.003)
$\hat{\delta}_\pi$	0.710***	0.808***	0.667***	0.779***	0.642***
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
$\hat{\delta}_x$	0.350*	0.345*	0.417**	0.679**	0.174
	(0.095)	(0.075)	(0.043)	(0.018)	(0.111)
R^2	0.378	0.390	0.376	0.391	0.406
J-test p-value	0.435	0.349	0.573	0.514	0.230

Notes: p-values in parentheses. */**/*** denote significance at the 10%/5%/1% level. GMM estimation with HAC weighting; instruments: a constant, the lagged change in import prices, a set of lagged macroeconomic indicators (inflation, unemployment, money growth, new manufacturing orders, Economic Sentiment Indicator, and GDP growth), and the lagged output gap. Aggregate and Weighted denote, respectively, $x_{t,S}^{A,UKL}$ (log difference of aggregated actual and potential output) and $x_{t,S}^{W,UKL}$ (firm-weighted average) microdata-based output gaps; boosted variants use the boosted-HP trend of Phillips and Shi (2021). GDP, HP is a conventional Hodrick–Prescott-filtered ($\lambda = 1600$) gap of aggregate Polish GDP. The dependent variable π_t is the quarterly consumer price index, corresponding quarter of the previous year = 100; the estimated intercepts and persistence coefficients therefore imply a long-run mean of about 101–102, i.e. annual inflation of roughly 1–2%. The J-test p-value is for the GMM overidentifying-restrictions test;

Three main findings emerge. First, the lagged-inflation coefficient $\hat{\delta}_\pi$ is statistically significant ($p < 0.001$) across all specifications, indicating strong persistence in Polish CPI inflation. Second, the output-gap coefficient $\hat{\delta}_x$ is positive in all specifications, consistent with the Phillips-curve relationship, although its statistical significance varies across output-gap measures. Both microdata-based measures are statistically significant: the aggregate output gap at the 10% level and the weighted output gap at the 5% level, under both the unboosted and boosted specifications. By contrast, the conventional macro HP-filtered GDP gap is not statistically significant ($p = 0.11$). Thus, the microdata-based measures provide stronger statistical evidence of a relationship between cyclical economic activity and inflation than the conventional macroeconomic benchmark, with the evidence particularly robust for the weighted output gap.

Third, the GMM overidentification test does not reject the validity of the instrument set in any specification, with p-values ranging from 0.23 to 0.57. The results therefore provide no evidence against the overidentifying restrictions across the five Phillips-curve specifications.

Overall, the results indicate that the quarterly microdata-based output gaps contain information about inflationary pressures in the Polish economy that is not captured to the same extent by the conventional macro HP-filtered GDP gap. In the GMM specifications with macroeconomic instruments, this relationship is particularly pronounced for the weighted microdata-based measure.

5 Conclusions

This paper develops a quarterly, firm-level, production-function-based output gap for Poland using two Statistics Poland microdata sources – the Quarterly Non-Financial Enterprises Survey and the Business Tendency Survey – over 2008Q1–2025Q1. The framework extends the annual approach of Błażej et al. (2025) to quarterly frequency by incorporating firm-level seasonal adjustment prior to trend extraction, the boosted Hodrick–Prescott filter of Phillips and Shi (2021), and formal business-cycle dating, including phase-duration analysis and business-cycle clocks. It also extends the sectoral coverage beyond manufacturing to construction and services and evaluates the real-time reliability of the resulting sectoral gaps through an expanding-sample vintage exercise. More generally, the approach differs from aggregate production-function methodologies in three important respects. Capacity utilisation enters the production function as a *directly observed* firm-level input, allowing measured productivity to be separated into utilisation-adjusted productivity and a capacity-utilisation component. Input elasticities are estimated using a control-function (Olley–Pakes/ACF) procedure separately for each of 62 two-digit NACE divisions rather than imposed from factor shares. Finally, potential output is constructed from the bottom up by extracting potential components from firm-level input series and subsequently aggregating firm-level potential output.

The resulting output gap captures the main cyclical episodes in the Polish economy since 2008 and identifies six complete contractions, lasting on average 6.2 quarters, compared with 5.4 quarters for completed expansions. Its decomposition provides additional information on the sources of these fluctuations. Capacity utilisation makes the largest negative contribution during the downturns of 2009, 2012–2013, and 2020, whereas the capital component is more prominent in 2015–2016, 2022–2023, and the pronounced downturn in 2024Q1. The measure also contains information relevant for inflation dynamics. In GMM Phillips-curve specifications with macroeconomic instruments, both microdata-based output gaps enter with the expected positive sign and are statistically significant, whereas the conventional Hodrick–Prescott gap of aggregate GDP is not. The main cyclical results are also robust to the choice between the two control-function specifications considered in the paper: replacing investment with intermediate inputs as the productivity proxy (LP-ACF) yields correlations above 0.97 with the baseline estimates and leaves the dating of turning points unchanged.

An important result of the analysis is the pronounced cross-sectoral heterogeneity in real-time reliability. Manufacturing exhibits relatively stable real-time estimates: its mean end-point revision is statistically indistinguishable from zero, and the real-time sign of the gap coincides with the benchmark in 88% of quarters. Construction and services display substantially greater instability. Both exhibit statistically significant systematic end-point revisions, with earlier vintages tending to place the gap above the level subse-

quently indicated by the benchmark. The problem is particularly pronounced for services. The real-time root-mean-square revision, 1.24 pp, amounts to 1.65 times the standard deviation of the services gap, while the real-time and benchmark estimates agree on its sign in only 25% of quarters. Consequently, for most quarters in 2023–2024, the real-time estimate indicates a positive services gap whereas the corresponding benchmark estimate is negative.

The services results also illustrate why the absolute magnitude of revisions alone is insufficient for assessing real-time reliability. Although absolute revisions are smaller for services than for construction, the services gap itself exhibits relatively limited cyclical variation, with a standard deviation of only 0.75 pp. Revisions of a few tenths of a percentage point can therefore be large relative to the underlying cyclical signal and may be sufficient to change the estimated sign of the gap. This suggests that real-time reliability should be evaluated relative to the magnitude of the cyclical variation in the estimated series, rather than solely in terms of absolute revision statistics. More broadly, the vintage exercise demonstrates the importance of evaluating microdata-based output-gap measures in real time: favourable full-sample properties do not necessarily imply equally reliable assessments at the end of the sample.

Several limitations should be considered when interpreting these results. First, the QNFES covers only firms with 50 or more employees, so the quarterly output gap primarily reflects cyclical conditions among medium-sized and large enterprises. Extending the framework to micro and small firms would require quarterly data with broader size coverage than are currently available. Second, matching QNFES with BTS reduces the number of firms available for estimation relative to either survey separately. Estimates for service divisions represented by relatively few firms, including some divisions with negative capital or capacity-utilisation elasticities, should therefore be interpreted with caution. In the economy-wide aggregation, however, their contribution is determined by their gross-value-added weights. Third, the survey-based concept of capacity utilisation may be less well suited to activities in which productive capacity is not closely related to physical capital. This may contribute to the greater dispersion of $\hat{\beta}_{u,d}$ across service divisions. Fourth, potential output necessarily depends on the method used to extract the trend component. Although the results are reported using conventional and boosted Hodrick–Prescott specifications, the vintage analysis shows that end-of-sample uncertainty remains important, particularly for construction and services.

These limitations suggest several directions for future research. A first priority is to examine whether the real-time reliability of the sectoral gaps can be improved through alternative trend-extraction methods, including sector-specific smoothing parameters, one-sided filters, or model-based approaches designed explicitly for real-time estimation. Services provide a particularly relevant setting for evaluating such alternatives given the magnitude of the revisions documented in the vintage exercise. Second, the firm-level

structure of the framework could be exploited more extensively. In addition to sectoral estimates, it can be used to construct regional gaps and gaps by firm size, while the cross-sectional dispersion of firm-level gaps may itself provide information about the cyclical state of the economy that is lost through aggregation in Section 2.3. Third, the quarterly output gap could be used as an input into the calculation of a cyclically adjusted fiscal balance, which would additionally require quarterly information on cyclically sensitive government revenues and expenditures. Finally, applying the framework to other economies with comparable firm-level production and capacity-utilisation data would provide an opportunity to assess whether the cross-sectoral differences in real-time reliability documented for Poland are also present in other institutional and economic settings.

Conflicts of interest

The authors declare no conflict of interest.

Data Availability Statement

The firm-level microdata underlying this study – the Quarterly Non-Financial Enterprises Survey (QNFES) and the Business Tendency Survey (BTS) – are confidential and were accessed under Statistics Poland's statistical confidentiality regime. The data are not publicly available and cannot be shared by the authors. Researchers seeking access to these or comparable confidential microdata may apply directly to Statistics Poland, subject to its data-access, data-protection, and statistical confidentiality procedures.

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

Basu, S. and Fernald, J. G. (2009). What do we know (and not know) about potential output? *Federal Reserve Bank of St. Louis Review*, 91(4):187–214.

Błażej, M., Górajski, M., and Ulrichs, M. (2025). Microdata-based output gap estimation using business tendency surveys. *Journal of Economic Dynamics and Control*, 174:105068.

Błażej, M., Górajski, M., and Ulrichs, M. (2025b). Synchronization and similarity between regional and sectoral output gaps in the Polish manufacturing industry. *Statistics in Transition new series*, 26(2):85–111.

Blondeau, F., Planas, C., and Rossi, A. (2021). *Output Gap Estimation Using the European Union's Commonly Agreed Methodology: Vade Mecum & Manual for EUCAM Software*, volume 148. European Commission, Directorate-General for Economic and Financial Affairs, Luxembourg.

Burns, A. F. and Mitchell, W. C. (1946). *Measuring Business Cycles*. National Bureau of Economic Research, New York.

Chalaux, T. and Guillemette, Y. (2019). The OECD potential output estimation methodology. *OECD Economics Department Working Papers*, No. 1563.

Christiano, L. J. and Fitzgerald, T. J. (2003). The band pass filter. *International Economic Review*, 44(2):435–465.

European Commission (2023). *The Joint Harmonised EU Programme of Business and Consumer Surveys: User Guide*. European Commission, Brussels.

Górajski, M. and Błażej, M. (2020). A control function approach to measuring the total factor productivity of enterprises in Poland. *Bank i Kredyt*, 51(3):293–316.

Havik, K., Morrow, K. M., Orlandi, F., Planas, C., Raciborski, R., Röger, W., Rossi, A., Thum-Thysen, A., and Vandermeulen, V. (2014). The production function methodology for calculating potential growth rates & output gaps. *European Economy - Economic Papers*, 535.

Hodrick, R. J. and Prescott, E. C. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16.

Hyndman, R. J., Koehler, A. B., Snyder, R. D., and Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18(3):439–454.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70(2):317–341.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.

Orphanides, A. and van Norden, S. (2002). The unreliability of output-gap estimates in real time. *Review of Economics and Statistics*, 84(4):569–583.

Phillips, P. C. B. and Shi, Z. (2021). Boosting: Why you can use the HP filter. *International Economic Review*, 62:521–570.

Rovigatti, G. and Mollisi, V. (2018). Theory and practice of total-factor productivity estimation: The control function approach using Stata. *The Stata Journal*, 18(3):618–662.

Statistics Poland (2023). *Business Tendency Survey Methodology*. Statistics Poland, Warsaw.

Tallman, E. W. and Zaman, S. (2017). Forecasting inflation: Phillips curve effects on services price measures. *International Journal of Forecasting*, 33(2):442–457.

A Firm-level variable definitions

Firm-level global output is defined as

$$output = revNetP + costsOwn + revNetCM - tax - valCM + \Delta goodsU + \Delta goodsF, \tag{21}$$

where $revNetP$ is net revenue from products, $costsOwn$ own production costs, $revNetCM$ net revenue from commodities, $valCM$ the value of sold commodities, and $\Delta goodsU$, $\Delta goodsF$ the change in semi-finished and finished goods, respectively. Intermediate consumption is

$$intermConsump = costsMatEnerg + costsExternalServ + costsOtherIC - costsBusTrip, \tag{22}$$

gross value added is $Y = output - intermConsump$, and capital is $capital = \overline{tangFixAss} + \overline{intangAss}$ (average tangible fixed and intangible assets). All value variables are deflated using two-digit NACE sector-specific deflators.

B GMM estimation of the production function

The second-stage GMM moment condition used to recover $\beta_{u,d}$, $\beta_{k,d}$, and $\beta_{l,d}$ is

$$y_{it} = g\Big(\hat{f}(u_{i,t-1}, k_{i,t-1}, l_{i,t-1}) - \beta_{u,d}u_{i,t-1} - \beta_{k,d}k_{i,t-1} - \beta_{l,d}l_{i,t-1}\Big) + \beta_{u,d}u_{it} + \beta_{k,d}k_{it} + \beta_{l,d}l_{it} + \varepsilon_{it} + \xi_{it}, \tag{23}$$

with instruments given by current capital k_{it}, lagged labour $l_{i,t-1}$, lagged capacity utilisation $u_{i,t-1}$, and lagged investment $i_{i,t-1}$. Standard errors are obtained by bootstrap.

C Robustness: OP-ACF versus LP-ACF production-function estimation

Section 2.1 estimates the sector-specific production function (1) using the Olley–Pakes control-function method with the Ackerberg–Caves–Frazer correction (OP-ACF), in which capital, labour, and capacity utilisation are approximated in the first stage by a polynomial in the firm's investment proxy. A widely used alternative proxy-variable estimator replaces investment with the firm's intermediate-input (materials and energy) outlays as a proxy for unobserved productivity – the Levinsohn–Petrin/ACF specification (LP-ACF) (Levinsohn and Petrin, 2003; Ackerberg et al., 2015). Because the two estimators rely on different identifying assumptions regarding the observable variable used to proxy for

unobserved productivity, comparing them provides a useful robustness check on whether the quarterly output gap depends materially on the specific control-function specification. We therefore re-estimate all steps of Sections 2.1–2.3 under the LP-ACF specification and recompute the economy-wide output gap, for both the aggregate measure $x_{t,S}^{A,UKL}$ (the log difference of aggregated actual and potential output) and the weighted measure $x_{t,S}^{W,UKL}$ (the firm-weighted average of firm-level gaps), over 2008Q1–2025Q1.

Aggregate output gap. Figure 9 compares the aggregate output-gap series obtained under the baseline OP-ACF and alternative LP-ACF estimators. The two approaches produce very similar output-gap paths. The correlation between the OP-ACF and LP-ACF aggregate gaps is 0.976, while the mean absolute difference between the two series is only 0.21 percentage points and the root-mean-square difference is 0.29 percentage points. Relative to a sample standard deviation of the aggregate output gap of 0.94 percentage points (Appendix D), these differences indicate that the choice of proxy variable accounts for only a small share of the overall variation in the estimated gap. The largest discrepancy occurs in 2009Q4, following the Global Financial Crisis trough (−0.99% under OP-ACF versus −0.11% under LP-ACF, a difference of 0.88 percentage points). Despite this difference, the two estimators yield the same sign and broadly similar magnitudes across the main cyclical episodes, differing primarily in the speed at which the estimated gap returns towards zero in subsequent quarters. There is also little evidence of a systematic level difference between the two estimates: the mean OP-ACF aggregate gap over the sample is −0.78%, compared with −0.66% under LP-ACF.

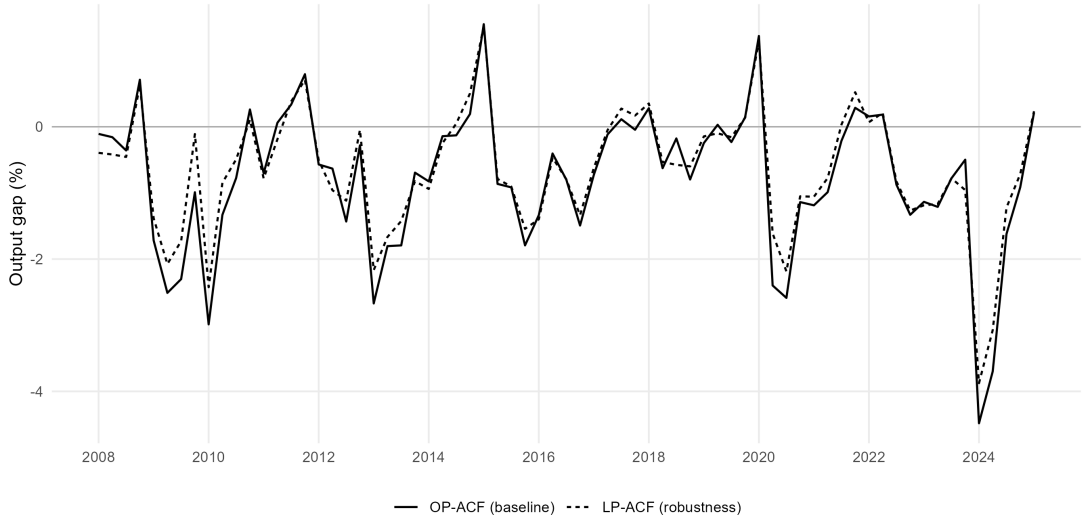


Figure 9: Aggregate quarterly microdata-based output gap $x_{t,S}^{A,UKL}$ under the OP-ACF (baseline, solid line) and LP-ACF (robustness, dashed line) production-function estimators, 2008Q1–2025Q1.

Weighted output gap. Figure 10 presents the analogous comparison for the weighted output gap $x_{t,S}^{W,UKL}$ (Appendix D). As for the aggregate measure, the OP-ACF and LP-ACF estimates are highly consistent. The correlation between the two series is 0.975, while the mean absolute difference is 0.16 percentage points and the root-mean-square difference is 0.22 percentage points, compared with a sample standard deviation of the weighted output gap of 0.78 percentage points. The largest discrepancy occurs in 2023Q4, during the post-energy-shock adjustment (-0.10% under OP-ACF versus -0.72% under LP-ACF, a difference of 0.62 percentage points). Discrepancies of a similar magnitude occur at the COVID-19 trough (2020Q2, 0.56 pp) and the Global Financial Crisis trough (2009Q3, 0.56 pp). There is also little evidence of a systematic level difference between the two estimates: the mean OP-ACF weighted gap is -0.39%, compared with -0.34% under LP-ACF. Overall, the two specifications yield very similar cyclical profiles, agreeing on the sign and timing of the main cyclical episodes.

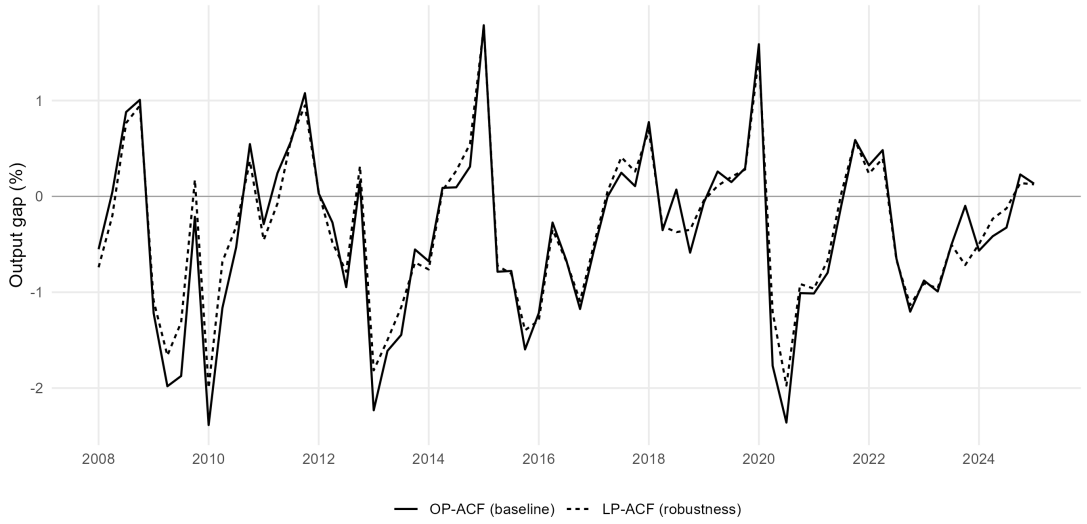


Figure 10: Weighted quarterly microdata-based output gap $x_{t,S}^{W,UKL}$ under the OP-ACF (baseline, solid line) and LP-ACF (robustness, dashed line) production-function estimators, 2008Q1–2025Q1.

For both the aggregate and the weighted measure, turning points – the peaks and troughs used for the business-cycle dating exercise in Section 2.4 – occur in the same quarters under both specifications. Overall, the close agreement between the OP-ACF and LP-ACF estimates indicates that the quarterly output gap, whether expressed as an aggregate or a weighted measure, is not sensitive to the choice between these two standard control-function specifications. The main cyclical patterns are therefore robust to replacing investment with intermediate inputs as the proxy for unobserved productivity.

D Weighted output-gap results

The main text (Section 4.2) reports results for the *aggregate* output gap $x_{t,S}^{A,UKL}$, defined as the log difference between aggregated actual and potential output. This appendix presents the corresponding results for the *weighted* output gap $x_{t,S}^{W,UKL}$, defined as the firm-weighted average of firm-level gaps (Equation (12)). The two measures need not coincide because the weighted measure gives each firm equal weight regardless of size, whereas the aggregate measure is effectively output-weighted and therefore more strongly influenced by large firms and divisions.

The two series are strongly, although not perfectly, correlated ($\rho = 0.82$ over 2008Q1–2025Q1). The aggregate gap is more volatile, with a standard deviation of 0.94 percentage points compared with 0.78 pp for the weighted gap, and more negative on average, with a sample mean of -0.78% compared with -0.39%. These differences are consistent with the greater influence of large firms and divisions on the aggregate measure. Figure 11 plots the weighted output gap together with its capital-, labour-, and capacity-utilisation-based decomposition, providing the weighted-measure counterpart to Figure 2 in the main text.

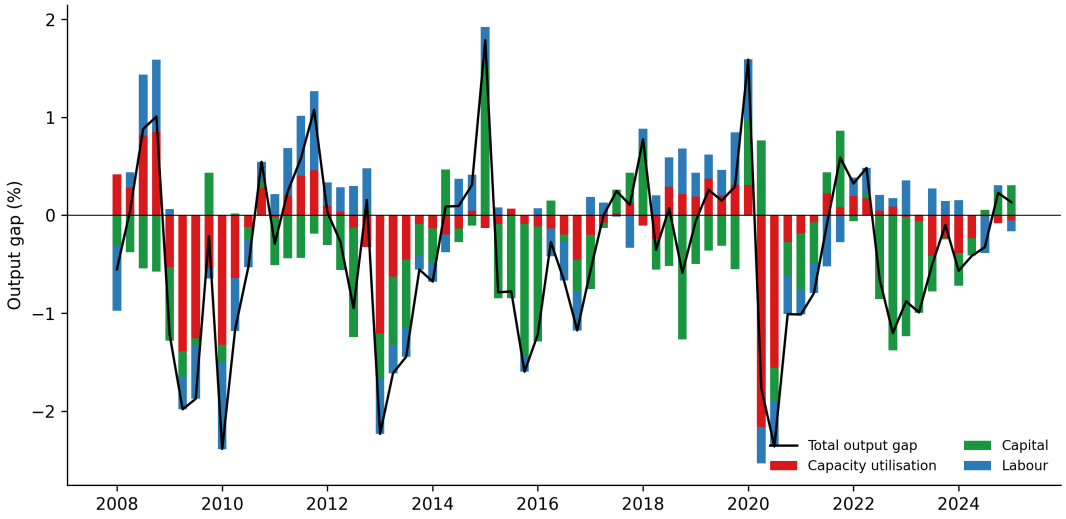


Figure 11: Weighted microdata-based output gap $x_{t,S}^{W,UKL}$ in Poland, 2008Q1–2025Q1, decomposed into capital-, labour-, and capacity-utilisation-based components.

The broad cyclical pattern closely mirrors that of the aggregate measure: the same episodes are visible at approximately the same dates, but with smaller fluctuations and without the pronounced 2024Q1 collapse observed in the aggregate series. The weighted gap reaches its minimum of -2.00% in 2010Q1, during the capacity-utilisation-led relapse following the Global Financial Crisis trough, and its sample maximum of $+1.77\%$ in 2015Q1, coinciding with the aggregate measure's peak but with a smaller capital contribution. At the troughs of most of the main contractions, the dominant component is the same under both measures: capacity utilisation during the Global Financial Crisis,

the 2012–2013 slowdown, and the COVID-19 shock, and capital during the 2015–2016 slowdown and the energy-price shock. The exception is the shallow 2017–2019 downturn, which is capital-led under the weighted measure but reflects approximately equal contributions from capital and capacity utilisation under the aggregate measure.

The most pronounced difference occurs in 2024Q1. The weighted measure does not reproduce the aggregate gap's sharp decline to −4.48%, with the corresponding weighted trough remaining substantially shallower. This difference indicates that the 2024 contraction was concentrated among firms and divisions with relatively large output weights, whose influence is greater in the aggregate measure than in the firm-weighted measure. Despite these differences in amplitude, the cyclical timing remains highly consistent: turning points used for business-cycle dating (Section 2.4) fall within one quarter of each other under the two measures in every identified episode.

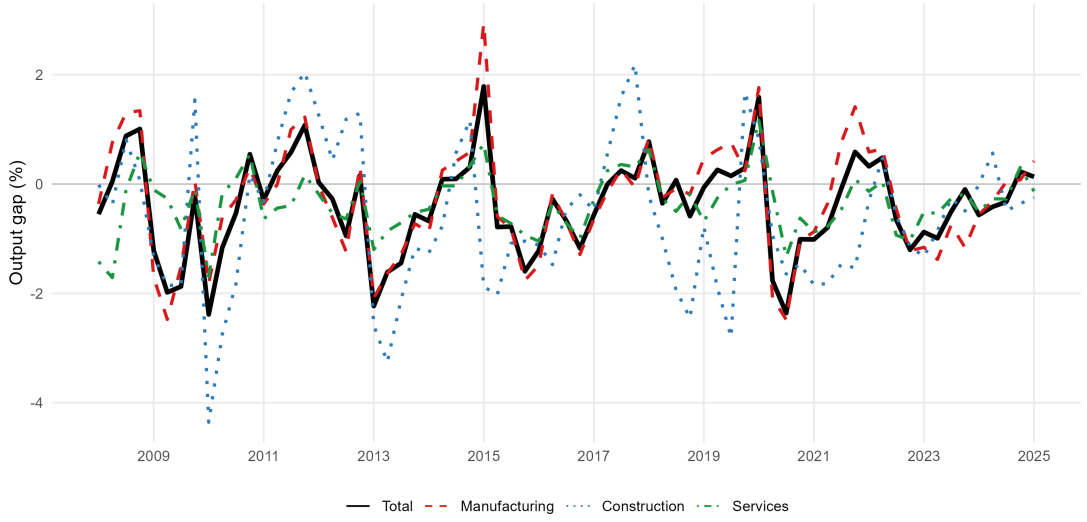


Figure 12: Weighted microdata-based output gap by sector: total, manufacturing, construction, and services, 2008Q1–2025Q1.

Figure 12 compares the quarterly weighted output gap for the whole economy with the corresponding weighted gaps for manufacturing, construction, and services. As under the aggregate measure, construction is the most volatile of the three sectors, with a standard deviation of 1.36 percentage points, compared with 1.05 pp for manufacturing and 0.55 pp for services. Construction's deepest single-quarter trough under the weighted measure, −4.36%, occurs in 2010Q1, and its peak, +2.18%, in 2017Q4 – both earlier than the corresponding aggregate-measure extrema, reflecting construction's different firm-weighting under the two measures. Manufacturing again tracks the weighted aggregate gap most closely, with a correlation of 0.96; it reaches its own sample peak of +2.94% in 2015Q1, coinciding with the weighted aggregate's sample maximum (+1.77%), and a COVID-19 trough of −2.48% in 2020Q3, close to the weighted aggregate's own trough for that episode (−1.97%). Services remains markedly less volatile (0.55 pp) and less synchronised with

the cycle: its correlation with the weighted aggregate gap is 0.78, with manufacturing 0.62, and with construction only 0.43. Construction's own correlation with the weighted aggregate gap is itself relatively modest, at 0.57, again pointing to a largely sector-specific, only loosely synchronised cycle. These patterns closely mirror those found for the aggregate measure, indicating that the relative volatility ranking and cross-sector correlation structure are robust to the choice between the aggregate and firm-weighted output-gap measures.

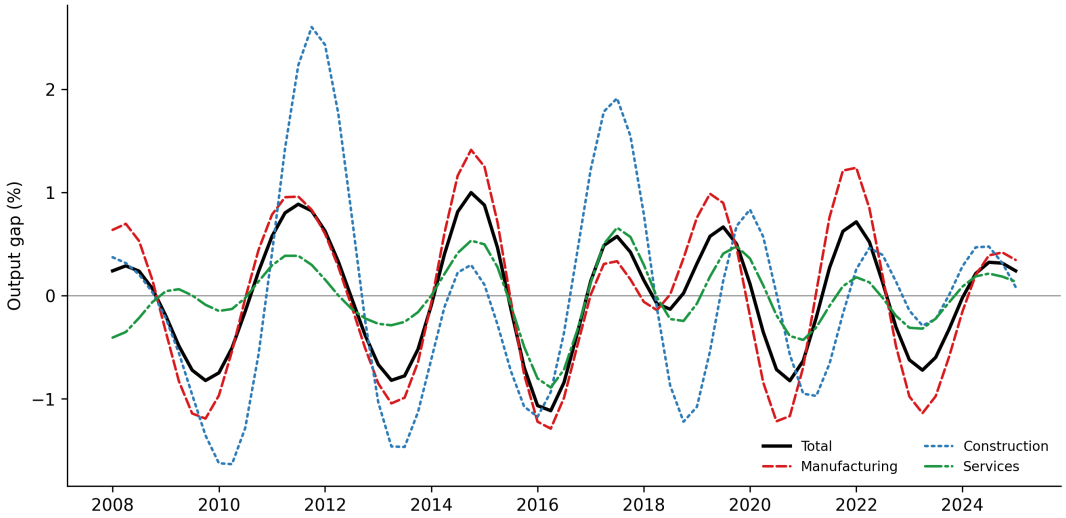


Figure 13: Cyclically adjusted weighted microdata-based output gap by sector: total, manufacturing, construction, and services.

Figure 13 reports the cyclically adjusted (Christiano–Fitzgerald filtered) weighted gaps, the counterpart to Figure 7 in the main text. The ordering of cyclical amplitudes is unchanged – construction is the most volatile sector and services the least – and the peaks and troughs of the weighted total coincide with, or fall within one quarter of, those identified on the aggregate measure in Table 6. The dating of the Polish business cycle is therefore not an artefact of the way firm-level gaps are aggregated.