ISSN (Print) 0473-453X
Discussion Paper No. 1178
ISSN (Online) 2435-0982

BUSINESS CYCLES WITH CYCLICAL RETURNS TO SCALE

Jay Hyun
Ryan Kim
Byoungchan Lee

June 2022

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Business Cycles with Cyclical Returns to Scale*

Jay Hyun[†] Ryan Kim[‡] Byoungchan Lee[§]

HEC Montréal Johns Hopkins University HKUST

October 21, 2021

Abstract

We study business cycles with cyclical returns to scale. Contrary to tightly parameterized production functions (Cobb-Douglas and Constant Elasticity of Substitution), we empirically identify strong input complementarity that leads to procyclical returns to scale. We therefore propose a flexible translog production function that allows complementarity-induced procyclical returns to scale, and we integrate this function into a standard medium-scale dynamic stochastic general equilibrium (DSGE) model. The estimated model with the procyclical returns to scale (i) features procyclical price markups, (ii) better matches the cyclicality of factor shares, and (iii) decreases by nearly half the contribution of markup shocks to output fluctuations.

***JEL* Codes**: C11, E23, E31, E32

Keywords: Business cycles, Translog production function, DSGE model, Returns to scale, Markup, Markup shocks

*We are grateful to David Andolfatto, David Berger, Mark Bils, Chris Conlon, Andres Drenik, Heitor S. Pellegrina (discussant), Katherine Ho, Yang Jiao, Loukas Karabarbounis, Amit Khandelwal, Jennifer La'O, Jae Won Lee, Andreas Mueller, Monica Morlacco (discussant), Seunghoon Na, Emi Nakamura, Hyunseung Oh, Zhesheng Qiu (discussant), Ricardo Reis, Bernard Salanie, Stephanie Schmitt-Grohé, Jón Steinsson, Martin Uríbe, Carlos Vegh, Jonathan Vogel, David Weinstein, Michael Woodford, Mark Wynne (discussant), Pierre Yared, Tack Yun, and the seminar participants at the Columbia colloquia, Asian Econometric Society Meeting, EconCon, EGSC, Midwest Macro Meeting, Tsinghua Workshop in Macroeconomics, WEAI, YES, KERIC, and HKUST-Jinan Macro Workshop for their helpful comments and suggestions. This paper subsumes previous papers circulated under the titles "Price-Cost Markup Cyclicality: New Evidence and Implications" and "Business Cycles with Input Complementarity". All errors are our own.

[†] jungsik.hyun@hec.ca, Jungsik Hyun. Department of Applied Economics, HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine Montréal, QC H3T 2A7, Canada

[‡] rkim59@jhu.edu, Sung Ryong Kim. Johns Hopkins University, School of Advanced International Studies, 1717 Massachusetts Ave NW, Office 749, Washington DC 20036

[§] bclee@ust.hk. Department of Economics, Hong Kong University of Science and Technology, 6068 Lee Shau Kee Business Building, Clear Water Bay, Kowloon, Hong Kong

1 Introduction

Standard business cycle models make strong *a priori* structural assumptions on the shape of the production function. The most widely used production function in macroeconomics is the Cobb-Douglas production function. Despite its convenient tractable features, this production function imposes an excessively restrictive structure on how firms substitute their inputs (elasticity of substitution), the productivity of each input (marginal product of input), and the productivity of all inputs together (returns to scale). This production function was often justified by the Kaldor (1957) growth facts, but the recent decline in the labor share (e.g., Karabarbounis and Neiman 2014) calls this justification into question. Many researchers acknowledge this limitation and have started to adopt a more general constant elasticity of substitution (CES) production function, but even this production function has important restrictions: one single constant parameter governs the elasticity of substitution among inputs, and returns to scale is typically assumed to be constant and fixed over time.

We empirically assess the plausibility of these restrictions by imposing and estimating a flexible translog production function (Christensen et al. 1973, 1975). Compared to a CES, the translog is another generalization of a Cobb-Douglas production function that allows more flexibility in input substitution, the marginal product of input, and returns to scale. Similar to the non-parametric production function estimation technique developed in the industrial organization literature (Gandhi et al. 2020), we utilize the first-order condition of firms to estimate the marginal product of input and to assess the variability in returns to scale. We employ standard panel data techniques with detailed industry-level panel data for the estimation.

Through our estimation, we find strong complementarity between labor and energy that leads to time-varying *procyclical* returns to scale. The idea of time-varying returns to scale is striking yet simple. It reflects the idea that when firms employ more factors during boom periods, there are synergies among these factors that lead to the larger aggregate marginal product of inputs and returns to scale compared to periods of recession. The procyclical movement in returns to scale also induces a procyclical wedge between the marginal product of input and the real input price, which is the price markup in standard macroeconomic models.

Motivated by our empirical evidence, we estimate a medium-scale dynamic stochastic general equilibrium (DSGE) model as in Smets and Wouters (2007), incorporating a flexible translog production function. Given the empirical importance of complementarity between labor and energy in generating procyclical returns to scale, we include energy input and allow a translog substitution parameter between labor and energy. Alongside the model with the translog production function, we estimate two other models with nested production functions: (i) a standard two-factor Cobb-Douglas production function with labor and capital which was originally used by Smets and Wouters (2007) and (ii) a three-factor Cobb-Douglas production function that additionally includes energy input.

The estimated model with the translog production function generates procyclical returns to scale and procyclical price markups, consistent with our empirical findings. We confirm significant complementarity between labor and energy, which induces procyclical returns to scale. In turn, the procyclical returns to scale leads to procyclical price markups in the otherwise standard DSGE model; the larger returns to scale in expansion decreases marginal costs and allows price markups to rise. To the contrary, the models with the Cobb-Douglas production functions—regardless of including or excluding energy input—generate countercyclical returns to scale and acyclical price markups, which are inconsistent with our empirical findings.[1]

Furthermore, we document that the model with a translog structure better matches the empirical cyclicality of input shares compared to other models with Cobb-Douglas production functions. In the data, labor shares are countercyclical, and energy shares are procyclical. However, under the Cobb-Douglas production function with fixed costs, all input shares are perfectly positively correlated and cannot match different cyclicality of input shares. We break this tight link between input shares and generate data-consistent labor and energy share cyclicality with the flexible translog production function. At the same time, our model generates procyclical profit shares, as in Smets and Wouters (2007) but for a different reason; our model relies on the procyclical price markups, whereas the their model relies on the countercyclical returns to scale.

Finally, in our model with the procyclical returns to scale, the contributions of price and wage markup shocks to output fluctuations are substantially smaller than those in the models with the Cobb-Douglas production function. As in Smets and Wouters (2007), we conduct the forecast error variance decomposition exercise for each of the three different models with the corresponding production functions. Although integrating the energy input into the conventional Cobb-Douglas production function has a negligible effect on the decomposition results, adding the translog structure with the procyclical returns to scale reduces nearly half of the contribution of the markup shocks on output. The comparisons of the Bayesian estimation results and the impulse responses to price and wage markup shocks across the different models reveal the importance of procyclical returns to scale in suppressing the markup shocks. The procyclical returns to scale changes Calvo parameters and amplifies responses of real variables, which in turn reduce the residual variations that are previously attributed to the price and wage markup shocks.

To the best of our knowledge, this paper is the first to investigate the role of procyclical returns to scale in a business cycle framework. Our paper is closely related to previous studies that go beyond constant returns to scale in business cycle analysis (Benhabib and Farmer 1994, 1996; Schmitt-Grohé

[1] We focus on the *unconditional* returns to scale and markup cyclicality when comparing the model predictions with the data. Although the Smets and Wouters (2007) model generates countercyclical markups conditional on demand shocks, it features acyclical markups unconditionally. In addition, the conventional model features countercyclical returns to scale because of the fixed cost of production. In our model with translog production, however, the countercyclical effects of the fixed costs on returns to scale are dominated by the procyclical effects of the input complementarity. See Section 3.3.1 for details.

2000), and a growing literature that generalizes an aggregate Cobb-Douglas production function (e.g., Antras (2004); Chrinko (2008); Karabarbounis and Neiman (2014); Cantore et al. (2015)).[2] Most previous studies reject a Cobb-Douglas production function and find complementarity among inputs beyond what Cobb-Douglas technology implies, similar to our analysis.[3] To integrate the general production function into the DSGE framework, we normalize the translog production function so that it preserves dimensionless parameters, as pioneered by De Jong (1967) and first incorporated in the DSGE framework by Cantore and Levine (2012) and Cantore et al. (2014). Regarding time-varying parameters, Koh and Santaeulàlia-Llopis (2017) propose a CES production function that features the time-varying elasticity of substitution. We complement previous studies by proposing a translog production function with the time-varying returns to scale.

The procyclical returns to scale speaks to the considerable literature that studies the countercyclicality of the price markup, which is a first-order building block in many sub-fields of macroeconomics.[4] Despite its importance, existing empirical evidence on price markup cyclicality is mixed.[5] This paper finds that integrating procyclical returns to scale leads to unconditional procyclical price markups in the standard medium-scale DSGE model. Our emphasis on the procyclical returns to scale differs from previous papers that emphasize other sources of procyclical price markup, such as wage rigidity (Nekarda and Ramey 2020), time-varying demand elasticity (Stroebel and Vavra 2019), and endogenous assortment (Anderson et al. 2020). The resulting procyclical price markup, in turn, decreases the contributions of markup shocks to output fluctuations compared to what is reported in Smets and Wouters (2007). These results imply that using a translog production function may alleviate the concerns raised in previous studies about the excessive importance of the markup shocks in new Keynesian models (see, e.g., Chari et al., 2009; Justiniano et al., 2010).

The remainder of this paper is structured as follows. Section 2 presents the industry-level data

[2]There are important studies that micro-found the aggregate production function with heterogeneous industry or firm models (e.g., Atalay 2017; Raval 2019; Oberfield and Raval 2021). In particular, Baqaee and Farhi (2021) show that the aggregate returns to scale can vary across time due to the change in allocative efficiency. Our paper instead extends the production function itself with the translog structure and infers the aggregate implications of using this more flexible production function.

[3]One notable exception is Karabarbounis and Neiman (2014), who find that labor and capital are substitutes; however, in estimating their parameters, they study long-term trends rather than business cycle movements. See, e.g., Hassler et al. (2019) for a discussion of different substitution patterns across short-run and long-run horizons among energy and other inputs.

[4]In the context of models of nominal rigidity, countercyclical markups conditional on demand changes are necessary to explain both procyclical wages and countercyclical unemployment (Rotemberg and Woodford 1991, Rotemberg 2013). In the study of monetary policy, many New Keynesian models suggest that central banks should target constant average markup for price stability (Goodfriend and King 1997). In the scholarship on price dynamics, countercyclical markups conditional on financial distortion explain missing disinflation during the Great Recession (Gilchrist et al. 2017). Ravn et al. (2006) find that introduction of deep habit formation substantially affects cyclicality of price-markups. Finally, Bils et al. (2018) find that unconditional countercyclical markups explain at least half of the cyclicality in the labor wedge.

[5]Some studies find that price markups are countercyclical, and other studies find that price markups are procyclical or acyclical. See, e.g., Bils (1987), Rotemberg and Woodford (1991), Rotemberg and Woodford (1999), Gali et al. (2007), Bils et al. (2013), Bils et al. (2018) for countercyclicality and Hall (2013), Nekarda and Ramey (2020), Stroebel and Vavra (2019), and Anderson et al. (2020) for procyclicality.

and the estimation results with a translog production function. Section 3 presents and estimates a medium-scale DSGE model with the translog production function and discusses the business cycle implications. Section 4 concludes.

2 Empirical Analysis

This section estimates production function coefficients under the translog production functional form assumption. We present the data used in this analysis and our empirical framework, which allows us to recover the marginal product of input, along with the estimation results. Then, we discuss the variability in returns to scale.

2.1 Data

This paper uses annual six-digit North American Industry Classification System (NAICS) industry-level data from the NBER-CES Manufacturing Industries Database. This database records detailed information on 473 manufacturing industries from 1958 to 2009. The information is compiled from the Annual Survey of Manufacturers and the Census of Manufacturers. The variables in this database include gross output (value of shipment), value added, and 5-factor inputs (production worker, non-production worker, capital, material, and energy) for each industry over time. These data also include industry-specific deflators for output, material, energy, investment, and wage bills for production workers and total employees. We report our summary statistics in Appendix A, and a more detailed explanation of this database can be found in Bartelsman et al. (2000).

The greatest advantage of the NBER-CES data over aggregate data is that they allow us to exploit both time-series and cross-sectional variations and corresponding panel data techniques to estimate production function parameters. A large variation in the data is especially important for our analysis, which seeks to relax strong functional form assumptions. The advantage comes with the cost, as our estimates come from the manufacturing sectors. In Appendix B.8, we use the Integrated Industry-Level Production Account (KLEMS) data that cover all sectors in the US economy for a shorter time period as a robustness check and confirm the strong complementarity between labor and energy input. The complementarity results are similar when we re-estimate the complementarity parameter with our DSGE model by using aggregate data that cover the entire US economy in Section 3.2.

2.2 Empirical Framework and Estimation Results

Estimaing a translog production function is challenging because it has an excessive number of parameters.[6] For example, for the five inputs available in the NBER-CES data, we must estimate twenty parameters with a translog production function, many more than the five parameters in a Cobb-Douglas production function and the six parameters in a CES production function. Even with the detailed industry-level data for many years, it is difficult to estimate all twenty parameters in the translog production function because of multicollinearity.[7]

To overcome the challenge of estimating many parameters, we exploit a firm's first-order condition in the spirit of the non-parametric identification method developed in the IO literature (Gandhi et al. 2020).[8] Firms' optimization conditions generate a relationship between the marginal product of each input and its price. Using this relationship and a panel data estimation technique, we recover the production part associated with the marginal product of input, which has significantly fewer parameters.

In estimating the first-order condition, we address potential estimation concerns by choosing the following three specifications. First, we use energy input to tightly link the marginal product of input (energy) from the real input (energy) price in the first-order condition. Second, we apply log-linearization and demean the variables to make the equations linear and to address potential endogeneity concerns. Third, we use lagged input prices as instrumental variables to address non-classical measurement errors and to rule out other remaining endogeneity issues. We conduct various robustness checks to address other potential concerns and report the results in Appendix B.

Estimation Framework For simplicity, consider the following translog production function with only two inputs, labor and capital:

$$\ln(Y) = \underbrace{\varepsilon^a + \beta_l \ln(L) + \beta_k \ln(K)}_{\text{Cobb-Douglas}} + \underbrace{\beta_{lk} \ln(L) \ln(K) + \frac{\beta_{ll}}{2} \ln(L) \ln(L) + \frac{\beta_{kk}}{2} \ln(K) \ln(K)}_{\text{second-order terms}}, \tag{2.1}$$

where Y is output, L is labor, K is capital, and ε^a is the log of total factor productivity. The first part of the production function is a conventional Cobb-Douglas function, which is a first-order approximation of a general production function. A translog production function extends this

[6]For example, Syverson (2011) wrote that "many researchers also use the translog form... is more flexible, though more demanding of the data". It is also difficult to calibrate parameters given that there is no previous work that integrates a translog production function into the business cycle model.

[7]For example, running an OLS regression of log output on 20 inputs with industry and time fixed effects makes more than half of the parameters statistically insignificant at the conventional level, including the labor input. Aside from other issues related to the estimation, such as the conditional mean independence assumption, the multicollinearity of inputs is likely to make the estimates statistically insignificant.

[8]The other way to proceed is to impose more structure in the estimation. In Section 3.2, we utilize the full structure of the business cycle model, estimate the key parameters, and confirm the industry-level empirical results.

approximation of a general production function to the second order. Assuming $\beta_{lk} = 0, \beta_{ll} = 0$, and $\beta_{kk} = 0$ recovers a Cobb-Douglas production function.

By generalizing Equation (2.1) to five different inputs available in the NBER-CES data, we have:

$$\ln(Y) = \underbrace{\varepsilon^a + \sum_i \beta_i \ln(V^i)}_{\text{Cobb-Douglas}} + \underbrace{\sum_i \sum_k \frac{\beta_{ik}}{2} \ln(V^i) \ln(V^k)}_{\text{second-order terms}} \text{ with } \beta_{ik} = \beta_{ki}, \tag{2.2}$$

where V denotes one of five different inputs indexed by i and k, namely, energy (e), production worker, non-production worker, capital, and material. Here, we allow for a flexible substitution structure among the five inputs. In contrast, conventional macroeconomic models use either two inputs (labor and capital) or three inputs (labor, capital, and material) with implicitly imposed restrictions on the substitution pattern among these inputs.

The simplest method to recover the parameters in Equation (2.2) is to regress log output on log inputs and treat the residual as the unobserved productivity. This approach has two key problems for our purpose. First, as we have already emphasized, it is extremely challenging to estimate all twenty parameters in this specification even with the rich variation available in the panel data. Second, as widely documented in the productivity estimation literature (e.g., Hall 1988; Evans 1992; Fernald 2014), flexible inputs are likely to be correlated with productivity, which generates inconsistent estimates of the parameters. For example, with the industry-level data, productive industries are likely to use more inputs relative to other industries.

To avoid these two concerns, we exploit a firm's first-order condition. Consider a firm's first-order condition with respect to an input V^i:

$$\underbrace{\frac{P^i}{P}}_{\text{real input price}} = \tau^i \underbrace{\left[\beta_i + \sum_k \beta_{ik} \ln(V^k)\right] \frac{Y}{V^i}}_{\text{marginal product of input}}, \tag{2.3}$$

where P^i is the nominal price of the input V^i, and τ^i is the wedge or gap between the real input price P^i/P and the marginal product of input V^i. The input-specific wedge τ^i allows Equation (2.3) to be consistent with a large class of models that deviates from the frictionless economy. Without frictions, the marginal product of input equals the real input price, and τ^i could be treated as a classical measurement error or the ex-post productivity shock as in Gandhi et al. (2020). Once researchers allow frictions, such as input adjustment cost (Hall 2004), imperfect competition in the output (Rotemberg and Woodford 1999) and input (Berger et al. 2019) markets, and financial friction (Jermann and Quadrini 2012; Arellano et al. 2019; Bigio and La'O 2020), such frictions generate a wedge between the marginal product of input and the real input price. This wedge is part of the labor wedge, which is the difference between the marginal product of labor and the marginal rate of substitution (see, e.g., Chari et al. 2007; Karabarbounis 2014; Bils et al. 2018). Note that assuming $\beta_{ik} = 0$ for all $i = 1, ..., 5$ recovers the conventional first-order condition under the Cobb-Douglas production function: $\frac{P^i}{P} = \tau^i \beta_i \frac{Y}{V^i}$.

To avoid the extra structural assumptions required to measure real output and the output price index (see, e.g., Hottman et al. 2016), we rewrite Equation (2.3) as follows:

$$s^i = \tau^i \left[\beta_i + \sum_k \beta_{ik} \ln(V^k) \right], \tag{2.4}$$

where $s^i = \frac{P^i V^i}{PY}$ is the input expenditure share out of total sales for the input V^i. The right-hand side is the wedge τ^i multiplied by an output elasticity with respect to the input V^i, which is a unit-free measure of the marginal product of input. The special cases of Equation (2.4) are used to calibrate parameters or inform price markup cyclicality in the previous macroeconomics models. Under the frictionless economy with Cobb-Douglas technology, an exponent of each input in the production function equals the corresponding input share: $s^i = \beta_i$. This restriction is often used to recover the Cobb-Douglas production function parameters from the observed income shares. In models of imperfect competition with the Cobb-Douglas production function, the wedge is interpreted as the inverse of the price-cost markup, and the inverse of input share identifies the markup up to a constant: $s^i = \beta_i \frac{1}{\text{markup}}$. This restriction allows researchers to inform on markup behavior with the input share.

To input the data into Equation (2.4), we log-linearize the equation around the steady state and allow the input share, wedge, and all inputs to vary across industry and time, which are indexed by industry j and time t, respectively:

$$\hat{s}^i_{jt} = \sum_k \delta_{ik} \hat{V}^k_{jt} + \hat{\tau}^i_{jt}, \tag{2.5}$$

where $\delta_{ik} \equiv \beta_{ik} \left(\frac{\bar{\tau}^i}{\bar{s}^i} \right)$, $\hat{x}$ denotes the log-deviation from the steady state value, and $\bar{x}$ denotes the steady state value for any variable x. The log-linearization eases the estimation by making the equation linear in parameters and is consistent with the DSGE analysis in Section 3.1.

We estimate Equation (2.5) in the panel data by double-demeaning across industry and time, which eliminates both the industry-specific and time-specific components. The empirical counterpart of Equation (2.5) is:

$$\hat{\hat{s}}^i_{jt} = \sum_k \delta_{ek} \hat{\hat{V}}^k_{jt} + \hat{\hat{\tau}}^i_{jt}, \tag{2.6}$$

where $\hat{\hat{x}}_{jt} = \ln x_{jt} - \frac{1}{J} \sum_{j=1}^{J} \ln x_{jt} - \left[\frac{1}{T} \sum_{t=1}^{T} \left(\ln x_{jt} - \frac{1}{J} \sum_{j=1}^{J} \ln x_{jt} \right) \right]$ for any variable x. Technically, the double-demeaning is identical to allowing the industry and time fixed effects; this method eliminates both the industry-specific and time-specific components, including the aggregate trend. Note that although we demean the variables and remove the aggregate components from all variables, our parameters of interest, δ_{ik}, match the aggregate parameters. Our benchmark estimation does not weight observations, but weighting different industries to improve the match with the aggregate moments does not change the main empirical results, as shown in Appendix B.6.

Since the input share s^i_{jt} and all five inputs V^k_{jt} are observed in the data, Equation (2.6) can be estimated by regressing the input share on all five inputs and treating the wedge $\hat{\tilde{\tau}}^i_{jt}$ as a residual. However, there are two major problems in estimating Equation (2.5) directly for every input i. First, the wedge term τ^i_{jt} may contain industry-time-varying components, such as an adjustment cost, that are correlated with inputs and generate a confounding relationship. For example, consider any positive aggregate shock that raises inputs in the production. If an industry j faces higher input adjustment costs relative to other industries, then industry j is likely to utilize fewer inputs relative to other industries.[9] Second, since the input share s^i contains the input V^i, there is a positive mechanical correlation between the input share $s^i \equiv V^i \times \frac{P^i}{PY}$ and the input V^i when V^i has a measurement error. See, for example, Berman et al. (2015) for the formal derivation of such a mechanical correlation.

To ease the first estimation concern, we choose an energy input as a choice variable and focus on estimating the energy efficiency (output elasticity with respect to energy). As a result of using energy input, there are fewer components in the wedge that can be correlated with V^k, particularly related to the adjustment cost. The energy input and the intermediates in general are known to have smaller adjustment costs relative to other inputs and are typically assumed away (e.g., Basu 1995; Bils et al. 2018). In addition, other potential concerns related to the monopsony power or heterogeneous input quality are mitigated when we focus on the energy shares.[10]

In addition to using the energy input, we utilize lagged double-demeaned input prices as instrumental variables to avoid the mechanical correlation and to relax concerns related to the remaining wedge term. As previously discussed, the input share and the input usage generate a positive mechanical correlation when the variables are measured with error. Unless a researcher has unusually detailed micro-level data, the input variables in any data have measurement error problems. For example, it is difficult to allow for bulk discounts or quality differences in material inputs or to

[9] Note that this problem resembles the issue in estimating Equation (2.2), which arises from the correlation between productivity and inputs.

[10] Regarding the monopsony power, there are fewer concerns on how firms exercise market power in the markets for energy inputs, and such friction would not appear as a wedge term. Previous studies have documented such frictions in labor input (e.g., Berger et al. 2019). In contrast, energy production in the U.S. has faced heavy regulation and other restrictions by Congress, such as tax preferences, spending subsidies, and environmental regulations. As a result, it is very unlikely that firms exert market power in their energy inputs. Regarding the heterogeneous input quality, energy input is likely to have homogeneous quality across industries relative to other inputs. Since input prices partially reflect the quality of input, higher input prices and input shares might reflect a higher quality of inputs, which will appear as a wedge in Equation (2.5). See, for example, De Loecker et al. (2016) for the structural treatment of input quality differences in the IO literature.

control education, experience, and specific skills of the labor input. Capital input is known to have a large measurement error even at the firm level (Collard-Wexler and De Loecker 2020), and the perpetual inventory method in the NBER-CES data requires an assumption on initial capital stock. Instrumenting inputs with the lagged input prices, which does not involve input usage, solves the mechanical correlation problem that arises from these measurement errors.

Regarding the remaining energy wedge $\hat{\hat{\tau}}^{e}_{jt}$, even after eliminating time- and industry-specific terms, there might be industry-time-varying wedge components that are correlated with inputs. By using the lagged input prices as instruments, we assume that the idiosyncratic $\hat{\hat{\tau}}^{e}_{jt}$ is not correlated with the idiosyncratic component of the predetermined input prices. If the idiosyncratic $\hat{\hat{\tau}}^{e}_{jt}$ is serially uncorrelated, then the lagged double-demeaned input prices satisfy the exogeneity assumption since they do not affect the current energy wedge. At the same time, the instrumental variables satisfy the relevance condition if they are autocorrelated and are correlated with input usage; Appendix A shows that these instruments are highly correlated with $\hat{\hat{V}}^{i}_{jt}$ for all i. In Appendix B.1, we also use two- and three-year lagged double-demeaned input prices and allow the correlation of $\hat{\hat{\tau}}^{e}_{jt}$ with idiosyncratic lagged input prices up to one year and find that the main empirical results remain robust.

In Appendices B.2 and B.4, we additionally conduct robustness exercises by putting specific interpretations in $\hat{\hat{\tau}}^{e}_{jt}$ based on previous macroeconomic models and find that controlling such elements in $\hat{\hat{\tau}}^{e}_{jt}$ has limited effects on the estimation results. For example, consider a multi-industry business cycle model with heterogeneous price rigidity across industries (e.g., Nakamura and Steinsson 2010). In this case, the energy wedge is price markup and can vary across time and industry. Moreover, if industries that face larger price rigidity alter their future markups (inverse wedge) and affect their input prices by changing the input usage relative to other industries, the exogeneity assumption of instruments could be violated. To address such a concern, we control the industry-time-varying measures of market power following previous studies, such as the price-cost markups (De Loecker et al. 2020) and the Lerner index (Gutierrez and Philippon 2017). We also include the measures of price rigidity and the inventory-to-sales ratio, which are known to be closely related to price markup, and measures of financial friction, adjustment costs, and fixed costs in production. Our main empirical results do not change with these alternative specifications, likely because double-demeaning at the detailed industry-time-level already eliminates most of the variations in the energy wedge that originates from the predictions of previous models.

Equation (2.6) clearly illustrates which variation in the data identifies the substitution pattern among energy and other inputs. Suppose that the coefficient of labor in Equation (2.6) is positive; the energy share increases with an increase in labor input, holding other inputs constant. Under the translog technology, such an increase in the share of energy is interpreted to be a result of an increase in energy efficiency that arises from an increase in labor input. In this case, the energy and labor inputs are complements, and the coefficient δ_{ek} captures the strength of the complementarity.

Note that a large magnitude of δ_{ek} translates to a much smaller magnitude of $\beta_{ek} = \delta_{ek}\frac{\bar{s}^e}{\overline{\tau}^e}$ due to the small energy share $\bar{s}^e$ in the data.

Since our empirical strategy is heavily motivated by the productivity estimation literature on industrial organizations, it is useful to compare our estimation technique with those in the literature. One key difference in our approach is that we rely on the representative production function, as in a typical DSGE model. Since the goal of our exercise is to discuss the aggregate parameters, we do not think that this assumption is particularly worrisome. Besides, in Section 3.1, we show that a model counterpart of the industry-level energy share (2.5) can be derived by aggregating the corresponding firm-level first-order condition with an assumption of the firm-level translog production function. Similarly, Appendix C.2 illustrates the representative translog production function by explicitly aggregating the firm-level translog production functions. Within the structure of our model, the interpretations of the production parameters are the same and the estimation assumptions are similar across different levels of aggregations.

Taking the industry-level production function as a given, our use of the first-order condition follows Gandhi et al. (2020), who estimate the first-order condition to allow a flexible substitution pattern in a production function. One advantage of using the industry-level data is the availability of the entity-level input price measures, which are rarely available in more micro-level data. We use input price deflators as instrumental variables to alleviate the mechanical correlation problems and the endogeneity concerns. In addition, we use them to measure the quantity of inputs at the entity-level and therefore avoid input price bias in using the product- or firm-level data (see, e.g., De Loecker and Goldberg 2014). Double-demeaning, instrumenting, and controlling variables to address the potential endogeneity of the energy wedge in estimating the first-order condition is similar to the methods that address the endogeneity of the total factor productivity in estimating the production function. Our use of instrumental variables is similar to the method of Doraszelski and Jaumandreu (2013), who use the first-order conditions and the lagged input prices to address the issue of simultaneity. The double-demeaning with instrumental variable approach resembles the dynamic panel data method (e.g., Arellano and Bond 1991; Blundell and Bond 1998), and controlling the potential energy wedge in our robustness exercise is similar to the control function approach (e.g., Olley and Pakes 1996; Levinsohn and Petrin 2003; Ackerberg et al. 2015). Note that for our DSGE analyses, we separately estimate the key complementarity parameter with the Bayesian method, which aligns with the estimate from this industry-level analysis.

Estimation Results Table 1 presents the estimated parameters in Equation (2.6). Column (1) is based on the baseline instrumental variable estimation with all five inputs. The most economically and statistically significant estimate is the coefficient in front of production workers, which measures the complementarity between production workers and energy input. The estimated parameter shows

Table 1: Estimation of Equation (2.6)

	Dependent Variable: Energy Share					
	IV (Lagged Input Prices)			OLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Energy	-0.557**	-0.399**	-0.603***	0.847***	0.842***	0.822***
	(0.230)	(0.161)	(0.229)	(0.029)	(0.029)	(0.032)
Production worker	1.694***	1.411***	1.775***	-0.328***	-0.326***	-0.467***
	(0.446)	(0.312)	(0.457)	(0.029)	(0.030)	(0.032)
Non-production worker	-0.593*	-0.454*	-0.560*	-0.119***	-0.124***	-0.224***
	(0.325)	(0.264)	(0.335)	(0.022)	(0.023)	(0.025)
Material	0.167	0.321**		-0.411***	-0.421***	
	(0.218)	(0.152)		(0.025)	(0.026)	
Capital	0.507		0.774**	-0.031		-0.243***
	(0.436)		(0.339)	(0.021)		(0.027)
J-test	6.04	6.76	5.8			
(p-value)	.3	.34	.45			
Observations	23220	23220	23220	23220	23220	23220

Note. Columns (1), (2) and (3) show the regression result with instrumental variables, and (4), (5), and (6) show the ordinary least square (OLS) result. For the instrumental variable analyses, we use both one-year lagged and two-year lagged input prices. For the implementation, we use the GMM specification with the weighting matrix that accounts for the arbitrary correlation among observations within sectors. Five different inputs are used in this regression, specifically, energy, material excluding energy, non-production workers, production workers, and capital. All inputs, lagged input prices, and the energy input share are logged and double-demeaned across industries and time. The standard errors in parentheses are clustered at the NAICS industry level. J-test and the p-value refer to Hansen's J-statistics and p-value for overidentifying restriction, respectively. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

that an increase in production workers of one percent leads to an increase in the energy share of approximately 1.7 percent. The coefficient in front of energy is negative, which shows that the energy input becomes less efficient as an economy uses more energy, thereby capturing the law of diminishing returns for energy. Non-production workers are a substitute, but the coefficient is marginally statistically significant. The J-test of Hansen (1982) does not statistically reject the validity of our instruments at the conventional level.

Since the coefficients of capital and material are not statistically significant in the baseline specification of column (1), we exclude each input and re-do the analysis in columns (2) and (3). We find that including these inputs separately leads to positive and statistically significant estimates, indicating that material and capital may be a complement to the energy input, like production workers. Given that the industry-level material and capital are highly correlated once they are double-demeaned, it is not surprising to observe that putting both variables together does not generate statistically significant estimates at the conventional level. Columns (4)-(6) show the OLS results. As expected, due to the mechanical correlation that arises from the measurement errors, the coefficient of energy input is positive, and it is difficult to make a meaningful inference.

Our results suggest a need for a more flexible production function with respect to energy input for macroeconomic models. Many of the coefficients in Equation (2.6) are economically and statistically different from zero, formally rejecting the Cobb-Douglas production function that requires the energy input share to be invariant with respect to any factor input. Appendix B.5 considers a CES production function, but the empirical results do not support this specification. In addition, for the robustness checks, Appendix B revisits the main empirical results by considering various other specifications, such as adjusting for the fixed costs in production, controlling more variables, allowing the regression weights, using a smaller number of inputs, and using KLEMS data. The key complementary between labor and energy, $\delta_{el} > 0$, is both economically and statistically significant at the conventional level in all the specifications; it ranges from 1.1 to 2.2 depending on the specifications, which are largely consistent with the Bayesian estimation results in Section 3.2.

2.3 The Returns to Scale Cyclicality

This section formally defines the returns to scale of the translog production function and accesses its cyclicality. Conceptually, returns to scale measures how much the percentage of output increases when all inputs increase by one percent. More precisely, by deriving the local elasticity of scale (Hanoch 1975; Epifani and Gancia 2006) for the translog production function $F(\{V^i\}; \varepsilon^a)$ in Equation (2.2), the industry-time-specific returns to scale is expressed as follows:

$$rts_{jt} = \sum_i \left[\beta_i + \sum_k \beta_{ik} \ln(V_{jt}^k) \right], \tag{2.7}$$

where $rts_{jt} \equiv \frac{\partial \log\left[F(\{\lambda V_{jt}^i\}; \varepsilon_t^a)\right]}{\partial \log(\lambda)}|_{\lambda=1}$ denotes the returns to scale, and λ reflects the proportional changes in all inputs. Under the conventional Cobb-Douglas production function, the returns to scale does not depend on inputs: $rts_{jt} = \sum_i \beta_i$. Under the translog production function, however, the returns to scale changes with the input usage. The degree of change depends on the parameter β_{ik}, which governs the substitution pattern among inputs. If inputs are complements (substitutes), an increase in one input raises (lowers) both the efficiency of other inputs and the returns to scale. The returns to scale in Equation (2.7) nests the constant returns to scale ($rts = 1$) embedded in the CES and the standard Cobb-Douglas production functions.

To examine the returns to scale with the estimated parameters, we follow the previous studies on returns to scale (Hall 1990; Basu and Fernald 1997) and additionally assume that the wedge does not differ across different inputs: $\tau_{jt}^i = \tau_{jt}$. This assumption still allows the common components of the wedge across different inputs, such as the price markup and fixed cost of production, and it is consistent with the standard DSGE models presented in Section 3. Despite its consistency with the previous work and the DSGE model, one potential concern about this assumption is the presence of an input-specific adjustment cost that is likely to differ across different inputs. In Appendix B.3, we explicitly integrate the input-specific adjustment cost and show that the returns to scale cyclicality results are robust to this concern.

With the common wedge assumption, we can rewrite the returns to scale by combining Equation (2.7) with the sum of Equation (2.4) across all input shares:

$$rts_{jt} = \tau_{jt}^{-1} \times s_{jt}^{\text{all}}, \tag{2.8}$$

where $s_{jt}^{\text{all}} \equiv \sum_i s_{jt}^i$ denotes the sum of all five input expenditure shares of total sales for industry j at time t. With an additional assumption that the wedge is the inverse price markup, one can link the returns to scale, price markup, and profit share, as in the previous studies of Hall (1990) and Basu and Fernald (1997): $rts_{jt} = \text{Markup}_{jt} \times (1 - \text{Profit Share}_{jt})$.

We log-linearize Equation (2.8) to assess the cyclicality of the returns to scale:

$$\widehat{rts}_{jt} = -\hat{\tau}_{jt} + \hat{s}_{jt}^{\text{all}}, \tag{2.9}$$

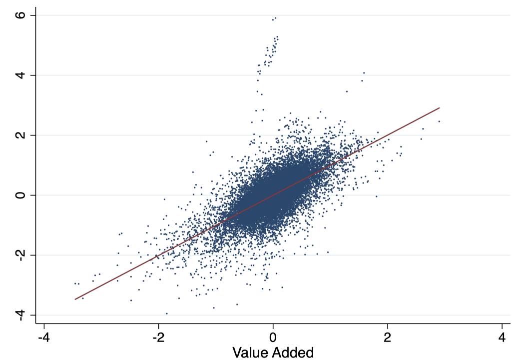
where $\hat{s}_{jt}^{\text{all}}$ is observed in the data and $\hat{\tau}_{jt}$ can be recovered as a residual of Equation (2.6) with the estimated parameters reported in Column (1) of Table 1; we use double-demeaned variables to recover the log-linearized variables. Due to the demeaning, our estimation strategy only identifies the wedge up to a constant and does not shed light on the *level* of returns to scale. However, we can still analyze the association between the returns to scale and the value added and infer the *cyclicality* of the returns to scale.

Figure 1 presents empirical evidence that the returns to scale is procyclical, and it arises from the procyclicality of the inverse wedge. Figure 1a shows a strong positive relationship between the returns to scale and the value added. We interpret this result as empirical evidence that supports the notion of the *procyclical* returns to scale; when industries experience larger (smaller) value added, they preserve larger (smaller) returns to scale. Correlating each component of the returns to scale and the value added reveals the importance of the inverse wedge in generating the procyclical returns to scale. The scatter plot of the inverse wedge and value added in Figure 1b closely replicates the scatter plot of the returns to scale and value added in Figure 1a, whereas the scatter plot of the sum of input shares (1 - profit share) and value added shows the negative relationship and does not contribute to the positive relationship plotted in Figure 1a.

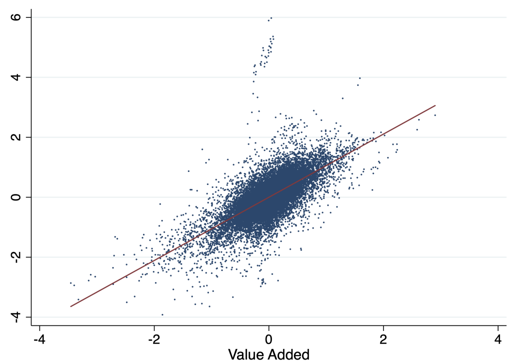
The economic intuition behind Figure 1 is straightforward once we specify a structural assumption

Figure 1: Returns to scale and its components

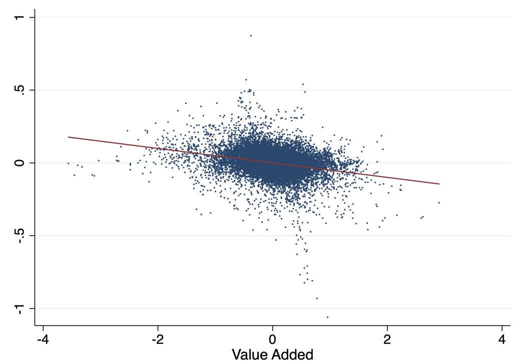
(a) Returns to Scale



(b) Inverse Wedge



(c) Sum of Input Shares



Note. Figure 1a assesses the variability in the returns to scale, and Figures 1c and 1b show the variability in the two components of the returns to scale, the inverse wedge and the sum of input shares. The y-axis is returns to scale, and it decomposes into the inverse wedge and the sum of input shares; the x-axis is value added. The returns to scale is recovered based on Equation (2.9). All the variables are in double-demeaned logged values, which shows how a 1% change in value added affects approximately x% of the returns to scale, inverse wedge, and sum of input shares. The slopes of the linear lines in Figures 1a, 1c, and 1b are 1, 1.05, and -.05, respectively.

that the wedge is the inverse price markup, which is consistent with the previous literature. Under this assumption, the strong input complementarity between energy and other inputs, such as production workers, leads to a larger increase in the marginal product of energy in expansion. As a result, the difference between the marginal product of energy and the real energy price—the price markup or the inverse marginal cost—becomes larger. The resulting decrease in the real marginal cost makes the production more efficient and increases the returns to scale. The rise in profit share decreases the economic efficiency in the expansion, but this effect is not strong enough to overturn the returns to scale procyclicality. Our results imply that there is a tight connection between the returns to scale cyclicality and price markup cyclicality, which we explore carefully in a standard medium-scale DSGE model in Section 3.

3 Macroeconomic Implications

Motivated by our empirical results, we explore the macroeconomic implications of the complementarity-induced procyclical returns to scale. We integrate the flexible translog production function into a standard medium-scale DSGE model (Smets and Wouters 2007). We re-estimate the model by using a Bayesian method with aggregate time-series data and confirm the procyclical returns to scale that arises from the complementarity between labor and energy. By comparing our benchmark model with the standard models of Cobb-Douglas production functions, we find that the model with the procyclical returns to scale (i) generates procyclical price markups, (ii) matches the different cyclicality of input shares, and (iii) decreases nearly half of the contribution of price and wage markup shocks to output fluctuation.

3.1 Model

This section describes how our model extends the DSGE model of Smets and Wouters (2007), which is nested in our model. Our discussion in this section focuses on illustrating the key differences from the standard model and the relationship with the empirical framework in Section 2. In particular, we characterize the translog production function with an energy input, the corresponding changes in the first-order conditions, and the modeling of the energy market. The other structure of the model follows Smets and Wouters (2007) closely. There are households, labor unions, final good producers, intermediate goods producers, the government and central bank, and global energy consumers and suppliers in the model. The model features sticky prices, sticky wages, utilization adjustment costs, investment adjustment costs, and consumption habits. The details of the model are relegated to Online Appendix A.

Production function Consider the following translog production function for intermediate good i at time t:

$$y_t(i) = \underbrace{\exp(\varepsilon_t^a)[k_t^s(i)]^{\beta_k}[l_t(i)]^{\beta_l}[e_t(i)]^{\beta_e}}_{\text{Cobb-Douglas}} \underbrace{\left(\frac{l_t(i)}{\bar{l}}\right)^{\beta_{el}\log(e_t/\bar{e})}\left(\frac{e_t(i)}{\bar{e}}\right)^{\beta_{el}\log(l_t/\bar{l})}}_{\text{second-order terms}} - \underbrace{\upsilon}_{\text{fixed costs}} \tag{3.1}$$

$$\text{with } \beta_k + \beta_l + \beta_e = 1,$$

where $k_t^s(i)$ is capital services used in production, $l_t(i)$ is labor, and $e_t(i)$ is energy. l_t and e_t are aggregate labor and energy, respectively, which individual firms take as a given when maximizing their profit, and $\bar{e}$ and $\bar{l}$ are steady-state values of e_t and l_t, respectively. Aggregate productivity $\exp(\varepsilon_t^a)$ follows an exogenous process, and υ is the fixed costs in production. The first part of the production function has a conventional Cobb-Douglas form with constant returns to scale, and the second part captures the second-order terms. All variables in Equation (3.1) are expressed with lowercase letters, which denote the detrended variables around the balanced growth path.[11]

We extend the standard two-factor Cobb-Douglas function with labor and capital to integrate the core of the empirical findings. Given the economically and statistically significant estimate, which indicates complementarity between labor and energy, we introduce an energy input $e_t(i)$ and allow the translog substitution structure between labor and energy via the substitution parameter β_{el} in the second-order terms. It is straightforward to see that our production function nests the three-factor Cobb-Douglas with $\beta_{el} = 0$ and the two-factor Cobb-Douglas used in Smets and Wouters (2007) with $\beta_{el} = 0$ and $\beta_e = 0$. We utilize these two special cases of our general translog production functions to emphasize how our new production function changes the traditional business cycle results.

By comparing the model's production function (3.1) with the empirical production function (2.2), we make two major changes to parsimoniously adapt the translog structure to the business cycle framework. First, we propose the three-factor translog production function with labor, capital, and energy; we neither additionally include material input with the input-output structure nor divide labor input into production worker and non-production worker. Although adding more inputs is a potentially interesting margin to explore, this is not necessary for the procyclical returns to scale results and could massively complicate the already complex analyses of the medium-scale DSGE model. Instead of adding more inputs to the model, in Appendix B.7, we redo our empirical analyses with a smaller number of factors by excluding materials and aggregating non-production and production workers into a single labor input. In all cases, we observe strong complementarity between labor and energy.[12] Because the empirical estimates are neither significant nor robust for the parameters other than the labor-energy complementarity, we focus on β_{el} and the corresponding translog structure in the model.[13] This minimal adjustment from the conventional framework highlights the role of the translog structure while avoiding overcomplications in the theoretical investigation.

Second, we normalize the production function. We introduce the inputs into second-order terms as a deviation from the steady-state values. This formulation carefully follows the previous studies that integrate the generalized production function into business cycle models to make the production function parameters dimensionless or unit-free (see, e.g., Cantore and Levine 2012; Koh and Santaeulàlia-Llopis 2017). By expressing the production function (3.1) as $\log \frac{y_t(i)+\upsilon}{\bar{y}+\upsilon}$, it is straightforward to see that the production function parameters are dimensionless with our normalization. Appendix C explicitly shows that the production parameters depend on units of

[11] Specifically, $y_t(i) = \frac{Y_t(i)}{\gamma^t}$, $e_t(i) = \frac{E_t(i)}{\gamma^t}$, $l_t(i) = L_t(i)$, and $k_t^s(i) = \frac{K_t^s(i)}{\gamma^t}$, where γ denotes the steady-state gross growth rate.

[12] In Appendix B.8, we consider the KLEMS data with three inputs. The strong complementarity between labor and energy remains robust.

[13] For example, the substitution parameter between energy and capital is not statistically significant based on the three-factor specification (Appendix B.7). The parameter for the energy square term is negative based on the NBER-CES data but positive based on the KLEMS data (Appendix B.8).

variables without making such a normalization. The normalization also makes the second-order term disappear at the steady-state and makes our production function directly comparable to the standard Cobb-Douglas specifications without changing the long-run balanced growth path.

When compared to the empirical specification (2.2), there are other minor changes in the normalized translog production function (3.1) that closely follows the specifications in Smets and Wouters (2007). We allow individual firm i's translog production function in the economy and introduce the second-order terms related to the returns to scale, l_t and e_t, as an aggregate externality for individual firms so that they do not choose their own returns to scale.[14] In addition, we incorporate the fixed costs in production υ as in the previous study. Given that our main empirical analyses in Section 2 do not allow these costs, Appendix B.4 revisits the empirical analyses and finds that the input complementarity and procyclical returns to scale and markup results are robust to the existence of the fixed costs. Finally, we include the labor-augmenting deterministic growth rate in the economy so that we can compare the model's outcome to the time-series data with trend. Appendix C presents the other general properties of the normalized translog production function (3.1).

First-order conditions Given real wage w_t, the real price of capital service r_t^k, and the real price of energy p_t^e, firm i solves its cost-minimization problem subject to the translog production function (3.1). The first-order conditions with respect to energy, labor, and capital are given by

$$p_t^e = mc_t(i)\frac{y_t(i)+\upsilon}{y_t(i)}\left(\beta_e + \beta_{el}\hat{l}_t\right)\frac{y_t(i)}{e_t(i)}, \tag{3.2}$$

$$w_t = mc_t(i)\frac{y_t(i)+\upsilon}{y_t(i)}\left(\beta_l + \beta_{el}\hat{e}_t\right)\frac{y_t(i)}{l_t(i)}, \tag{3.3}$$

$$r_t^k = mc_t(i)\frac{y_t(i)+\upsilon}{y_t(i)}\beta_k\frac{y_t(i)}{k_t^s(i)}, \tag{3.4}$$

where $mc_t(i)$ is the real marginal cost of production or the inverse price markup that arises from monopolistic competition, and $\hat{x}$ is the log deviation from the steady state value $\bar{x}$ for any variable x. As in our empirical analyses (Equation (2.3)), Equation (3.2) shows that the real energy price equals the marginal product of energy multiplied by the wedge that includes the real marginal costs and the term related to the fixed cost. Although we evaluate Equation (3.2) at the aggregate level in this model, evaluating the same equation at the industry-level recovers the industry-time-varying first-order condition used in Section 2.[15]

[14] This specification is similar to the externality assumption in the increasing returns to scale literature (e.g., Baxter and King 1991) or the redistributive shock introduced in Rios-Rull and Santaeulalia-Llopis (2010).

[15] Specifically, consider a continuum of firms, which is indexed by i and operates in industry j. $y_t(i,j)$ is given

Compared to the model with a Cobb-Douglas specification with an energy input, the only extensions with the translog are $\beta_{el}\hat{l}_t$ and $\beta_{el}\hat{e}_t$ in the first-order conditions (3.2) and (3.3), up to log-linearization. Clearly, β_{el} has a first-order effect on factor demand. Assuming the complementarity between labor and energy ($\beta_{el} > 0$), an increase in labor or energy input in expansion raises the marginal product of the other input more than the standard case. In contrast, there is a smaller increase in the marginal product of input if labor and energy are substitutes ($\beta_{el} < 0$). Note that changing β_{el} has no direct first-order effect in the production function (3.1). Our production function is identical to the Cobb-Douglas specification with log-linearization, as explicitly shown in Appendix C. This feature of the model is desirable given that our empirical framework in Section 2 relies solely on the first-order conditions to identify the complementarities between labor and energy inputs, which is consistent with our model. Within the structure of the standard model, we reassess our empirical results by re-estimating the key parameter β_{el} with the aggregate data.

Energy market Given that we introduce an energy input into the production function, we need to specify the energy market. We introduce a global energy market into our model where the real energy price p_t^e is determined subject to energy demand and supply shocks. We impose a parsimonious structure on the energy market to focus on the general production function and minimize deviations from the benchmark Smets and Wouters (2007) model.

Denote the global energy demand, excluding the US industrial usage e_t, as e_t^d. The energy market clearing condition is given by

$$e_t + e_t^d = e_t^s, \tag{3.5}$$

where e_t^s is the global energy supply. The global energy supply is determined by the energy price and exogenous disturbances:

$$\frac{e_t^s}{\bar{e}^s} = \left(\frac{p_t^e}{\bar{p}^e}\right)^{\kappa_s} \exp(\varepsilon_t^{es}), \tag{3.6}$$

where κ_s denotes the price elasticity of e_t^s, and the exogenous supply disturbances ε_t^{es} follow the AR(1) process of $\varepsilon_t^{es} = \rho_{es}\varepsilon_{t-1}^{es} + \eta_t^s$, where $\eta_t^s \sim (0, \sigma_{es}^2)$. We assume that ϕ_e fraction of e_t^s is produced domestically. Thus, the US net energy import is given by $e_t - \phi_e e_t^s$, which implies that the gross

by $f(k_t^s(i,j), l_t(i,j), e_t(i,j); l_t(j), e_t(j), \varepsilon_t^a)$, where $l_t(j)$ and $e_t(j)$ are aggregated labor and energy at the industry level, respectively, and $f(k_t^s(i), l_t(i), e_t(i); l_t, e_t, \varepsilon_t^a)$ denotes the production function $y_t(i)$ in Equation (3.1). Each firm (i,j) takes $l_t(j)$ and $e_t(j)$ as a given. Under this condition, the firm-level cost minimization problem implies that $s_t^e(i,j) = \tau_t(i,j)\left(\beta_e + \beta_{el}\hat{l}_t(j)\right)$, where $s_t^e(i,j)$ and $\tau_t(i,j)$ are the energy share and the wedge at the firm-industry-time-level, respectively. When firms in the same industry j are exposed to the same industry-level realization of Calvo shocks, these firms become symmetric, and we obtain $s_t^e(j) = \tau_t(j)\left(\beta_e + \beta_{el}\hat{l}_t(j)\right)$, which features industry-level labor in the marginal product of energy. Furthermore, cross-sectional aggregation implies that, up to log-linearization, $s_t^e = \tau_t\left(\beta_e + \beta_{el}\hat{l}_t\right)$, which is the same aggregate equation that we can derive from Equation (3.2).

domestic product (GDP) is $y_t^{GDP} = y_t - (e_t - \phi_e e_t^s)$.

The global energy demand, excluding US industrial usage, depends on the energy price p_t^e, exogenous disturbances to the demand ε_t^{ed}, US GDP y_{t-1}^{GDP}, and real interest rates $\mathbb{E}_{t-1}[R_{t-1}/\Pi_t]$, where R_t and Π_t represent gross nominal interest rates and gross inflation, respectively:

$$\frac{e_t^d}{\bar{e}^d} = \left(\frac{y_{t-1}^{GDP}}{\bar{y}^{GDP}}\right)^{\rho_{ey}} \left(\frac{\mathbb{E}_{t-1}[R_{t-1}/\Pi_t]}{\bar{R}/\bar{\Pi}}\right)^{\rho_{err}} \left(\frac{p_t^e}{\bar{p}^e}\right)^{-\kappa_d} \exp(\varepsilon_t^{ed}), \tag{3.7}$$

where κ_d denotes the price elasticity of the global energy demand. The exogenous demand disturbances ε_t^{ed} follow the AR(1) process of $\varepsilon_t^{ed} = \rho_{ed}\varepsilon_{t-1}^{ed} + \eta_t^d$, where $\eta_t^d \sim (0, \sigma_{ed}^2)$. Clearly, global economic activity positively affects energy demand (Balke and Brown, 2018; Kilian, 2009). We include lagged US GDP and real interest rates on the right-hand side as proxies for global economic activity. Furthermore, real interest rates capture the states of financial markets that might affect energy prices, as discussed in Basak and Pavlova (2016) and Kilian (2014).

The remaining parts of the model are identical to those in the Smets and Wouters (2007) model. See Online Appendix A for the full structure of the model.

3.2 Bayesian Estimation

This section discusses the Bayesian estimation by illustrating the construction of the likelihood function and by presenting the prior and the posterior. For the objective comparisons across models, we apply the Bayesian techniques, prior, and data identical to those in Smets and Wouters (2007) to both our benchmark model with the three-factor translog function and the model with the three-factor Cobb-Douglas function ($\beta_{el} = 0$). We also bring in the energy data and make relevant prior assumptions to estimate the new parameters introduced with the energy market in both models. The corresponding estimation results not only confirm the input complementarity ($\beta_{el} > 0$) we find in Section 2 but also reveal more rigid prices and wages and less persistent markup shocks. These changes are important in understanding the sources of the business cycle, which we analyze in Section 3.3.3.

We use a Kalman filter to compute the likelihood of a state-space system. For the state equation, we employ the algorithm suggested in Sims (2002). Our observation equation consists of nine variables, including the growth rate of the global energy supply $\Delta\log(E_t^s) \equiv \log(E_t^s) - \log(E_{t-1}^s)$, the logarithm of the real energy price $\log p_t^e$, and the seven major macroeconomic variables used in Smets and Wouters (2007, Equation (15)):

$$\begin{pmatrix} dlGDP_t \\ dlCONSUMPTION_t \\ dlINVESTMENT_t \\ dlWAGE_t \\ lHOURS_t \\ dlPRICE_t \\ FEDFUNDS_t \\ lENERGYPRICE_t \\ dlENERGY_t \end{pmatrix} = \begin{pmatrix} \bar{\gamma} \\ \bar{\gamma} \\ \bar{\gamma} \\ \bar{\gamma} \\ \bar{l} \\ \bar{\pi} \\ \bar{r} \\ \bar{p}^e \\ \bar{\gamma} \end{pmatrix} + \begin{pmatrix} \hat{y}_t^{GDP} - \hat{y}_{t-1}^{GDP} \\ \hat{c}_t - \hat{c}_{t-1} \\ \hat{i}_t - \hat{i}_{t-1} \\ \hat{w}_t - \hat{w}_{t-1} \\ \hat{l}_t \\ \hat{\pi}_t \\ \hat{r}_t \\ \hat{p}_t^e \\ \hat{e}_t^s - \hat{e}_{t-1}^s \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \nu_t \end{pmatrix}, \tag{3.8}$$

where $\hat{c}_t$, $\hat{i}_t$, $\hat{\pi}_t$, and $\hat{r}_t$ denote the log deviations of (detrended) consumption, investment, gross price inflation, and gross nominal risk-free return, respectively.

We obtain the global energy quantity data $\{E_t^s\}$ from BP Energy (2020). The real energy price is computed by the ratio of the producer price index for total energy to the GDP deflator. Because it is an index, we demean the logarithm of the energy price in Equation (3.8). For the other macroeconomic variables, we use the dataset constructed by Smets and Wouters (2007). Our sample period is from 1966:q1 to 2004:q4, which is identical to that in Smets and Wouters (2007). Because E_t^s data are available only at an annual frequency, we interpolate the annual series to construct a quarterly measure and introduce a measurement error ν_t in the observation equation. The standard deviation of ν_t is denoted by σ_ν. In total, we have nine observable variables, nine structural shocks (global energy supply and demand shocks and seven macroeconomic shocks in Smets and Wouters (2007)), and one measurement error. See Online Appendix B for further details of the data.

Table 2 presents the prior and the posterior of the parameters that arise from the integration of the translog structure and the energy market; we assume the same prior in Smets and Wouters (2007) for the common parameters used in the other two models. Among the new parameters, $s_e = \frac{\bar{e}}{\bar{e}^s}$ cannot be separately identified from the other parameters that characterize the dynamics of the energy market. Thus, we set s_e at the average value of $\frac{E_t}{E_t^s}$ in the data during our sample period, which is 0.08. Based on previous studies on the energy shares of value added, we assume that β_e has a Gamma distribution with a 5% mean and 2% standard deviation.[16]

For the prior of the substitution parameter β_{el}, we utilize our empirical estimate. By log-linearizing Equation (3.2) and taking cross-sectional average, we can relate β_{el} to δ_{el} in Equation

[16] The energy shares of value added are assumed to be 10% in Backus and Crucini (2000), 5.17% in Dhawan and Jeske (2008), 4.3% in Finn (2000), and 4% in Rotemberg and Woodford (1996).

Table 2: The prior and the posterior of the new parameters

Parameter	Prior			HKL		HKL-CD	
	Mean	St. Dev	Family	Posterior Mode	Credible Set (5%, 95%)	Posterior Mode	Credible Set (5%, 95%)
β_e	0.05	0.02	Gamma	0.023	(0.017, 0.033)	0.019	(0.013, 0.028)
β_{el}	0.1	0.025	Gamma	0.042	(0.031, 0.057)	-	-
ρ_{ey}	1	0.8	Normal	0.08	(0.02, 0.21)	0.12	(0.03, 0.26)
ρ_{err}	1	0.8	Normal	0.37	(0.16, 0.90)	0.37	(0.16, 0.91)
κ_d	0.1	0.08	Gamma	0.014	(0.0005, 0.167)	0.017	(0.005, 0.154)
κ_s	0.1	0.08	Gamma	0.12	(0.04, 0.16)	0.11	(0.04, 0.14)
σ_{ed}	0.8	2	Inv. Gamma	0.74	(0.66, 1.40)	0.77	(0.70, 1.37)
σ_{es}	0.8	2	Inv. Gamma	0.82	(0.56, 1.02)	0.78	(0.55, 0.91)
ρ_{ed}	0.9	0.05	Beta	0.9993	(0.9989, 0.9994)	0.9994	(0.9991, 0.9995)
ρ_{es}	0.9	0.05	Beta	0.9972	(0.9961, 0.9978)	0.9983	(0.9981, 0.9985)
$\bar{p}^e$	0	2	Normal	-0.31	(-3.51, 3.03)	-0.10	(-3.28, 3.30)
σ_ν	0.1	0.1	Inv. Gamma	0.05	(0.04, 0.24)	0.05	(0.04, 0.26)
ϕ_e	0.0337	0.0147	Gamma	0.033	(0.016, 0.075)	0.029	(0.015, 0.065)

Notes: HKL denotes our benchmark model with the translog production function. HKL-CD refers to the Cobb-Douglas specification with energy but without complementarity ($\beta_{el} = 0$).

(2.6) within the structure of this model by using the following equation:

$$\beta_{el} = [\delta_{el} + (\Phi - 1)\beta_l] \frac{\beta_e}{1 - \beta_e}, \tag{3.9}$$

where Φ is the gross price markup at the steady state.[17] When $\delta_{el} = 1.694$, which is our benchmark estimate in Table 1, and Φ, β_l, and β_e equal their prior means of 1.25, 0.65, and 0.05, respectively, Equation (3.9) implies that $\beta_{el} = 0.098$. When we multiply the standard error of $\hat{\delta}_{el}$ in Table 1 by $\frac{\beta_e}{1-\beta_e}$, which is evaluated at the prior mean of $\beta_e = 0.05$, we obtain 0.024. Thus, we assume that β_{el} has a Gamma distribution with a mean of 0.1 and standard deviation of 0.025; our posterior estimate and the corresponding theoretical results are nearly identical when we consider the alternative values of δ_{el} estimated in Appendix B.

For the rest of the parameters, we use the following priors. We demean $\log p_t^e$ in our observation equation and thus assume that $\bar{p}^e$ has a normal distribution with a zero mean and standard deviation of two. This choice is similar to the prior of $\bar{l}$ in Smets and Wouters (2007). For the standard deviation of the measurement error, we choose mean and standard deviation of 0.1. We decide the prior of the global share of the US energy production ϕ_e given the following considerations. The US net import of energy, E_t^{ni}, is given by $E_t - E_t^s \phi_e$. Then, $\phi_e = \overline{\left(\frac{E_t}{E_t^s}\right)} - \overline{\left(\frac{E_t^{ni}}{E_t^s}\right)} = s_e - \overline{\left(\frac{E_t^{ni}}{E_t^s}\right)}$. By subtracting the sample average of E_t^{ni}/E_t^s from s_e, we obtain the prior mean of ϕ_e, which is 0.0337. The sample

[17] Equations (A.70) and (A.80) in Online Appendix A are also used in the derivation.

Table 3: The Calvo and the markup shock parameters in the three models

	Parameter	Prior			Posterior Mode		
		Mean	St. Dev	Family	HKL	HKL-CD	S&W
ξ_p	Calvo sticky price	0.5	0.1	Beta	0.75	0.69	0.67
ξ_w	Calvo sticky wage	0.5	0.1	Beta	0.84	0.78	0.76
ρ_p	Price markup shocks:	0.5	0.2	Beta	0.83	0.91	0.89
μ_p	$\hat{\lambda}_t^p = \rho_p \hat{\lambda}_{t-1}^p + \eta_t^p - \mu_p \eta_{t-1}^p$,	0.5	0.2	Beta	0.69	0.77	0.73
σ_p	where $\eta_t^p \sim N(0, \sigma_p^2)$	0.1	2	Inv. Gamma	0.15	0.14	0.14
ρ_w	Wage markup shocks:	0.5	0.2	Beta	0.92	0.96	0.97
μ_w	$\hat{\lambda}_t^w = \rho_w \hat{\lambda}_{t-1}^w + \eta_t^w - \mu_w \eta_{t-1}^w$,	0.5	0.2	Beta	0.85	0.90	0.91
σ_w	where $\eta_t^w \sim N(0, \sigma_w^2)$	0.1	2	Inv. Gamma	0.23	0.24	0.25

Notes: HKL denotes our benchmark model with the translog production function. HKL-CD refers to the Cobb-Douglas specification with energy but without complementarity ($\beta_{el} = 0$). S&W is the Smets and Wouters (2007) model.

standard deviation of E_t^{ni}/E_t^s in the data is 0.0147, which is used as a prior standard deviation. For the shock processes, we rely on a preliminary regression analysis. Specifically, we estimate $\log E_t^s = c + dt + \kappa^s \log p_t^e + \varepsilon_t^{es}$ and a similar equation for $\log E_t^d$ by using the lagged values of Romer and Romer (2004) monetary policy shocks as instrumental variables. Based on the point estimates and standard errors, we set the prior mean and standard deviations of $\rho_{ey}, \rho_{err}, \kappa_d, \kappa_s, \sigma_{ed}, \sigma_{es}, \rho_{ed}$, and ρ_{es}. See Online Appendix B.2 for these regression results.

The posterior is computed by using a random walk Metropolis-Hastings algorithm with a chain length of 500,000. We consider three different models: our benchmark model (HKL), the model without input complementarity in production ($\beta_{el} = 0$; HKL-CD), and the Smets and Wouters (2007) model ($\beta_{el} = 0, \beta_e = 0$; S&W). The acceptance rates of the chains are 29%, 35%, and 23%, respectively.

As shown in Table 2, the following parameters are newly introduced and estimated compared to Smets and Wouters (2007). The energy share in the steady-state β_e is estimated to be 2.3%. The posterior mode of β_{el} is 0.042. If we use Equation (3.9) to compute the value of δ_{el} that corresponds to the posterior mode of β_{el}, Φ, β_l, and β_e, we obtain 1.36, which is within the range of one standard error of $\hat{\delta}_{el}$ in Table 1. The price elasticities of global energy demand and supply, κ_d and κ_s, respectively, are small at the posterior mode, which is consistent with the results in Hamilton (2009), Kilian (2009), and Kilian and Murphy (2012). Both energy supply and demand shocks are highly persistent. Furthermore, the sizes of these shocks are estimated to be larger than the other structural shocks because energy prices substantially fluctuated during the sample period. Finally, the measurement errors induced by the interpolation are estimated to be relatively small.

Among the other parameters common to the Smets and Wouters (2007) model, Table 3 reports the parameters that show the most noticeable changes across the three models: the parameters associated with the Calvo stickiness and the markup shocks. The estimation results show that our benchmark model features larger price and wage rigidity and less persistent markup shocks compared to the two other models with Cobb-Douglas production functions.[18] More rigid prices are intuitive, given the input complementarity ($\beta_{el} > 0$) reported in Table 2. As will be shown explicitly in Section 3.3.1, the input complementarity generates large countercyclical fluctuations in the marginal costs. As a result, the Bayesian estimator in our benchmark model prefers the stickier prices or the flatter price Philips curve to match the correlation of inflation and GDP observed in the data. Given the stickier prices, the stickier wages in our benchmark model are also intuitive. As the Bayesian estimator matches the real wage fluctuation in the data (Equation (3.8)), it changes the wage Philips curve parallel to the flatter price Philips curve and makes nominal wages more rigid.[19] Corresponding to the flatter Philips curves, the real variables' responses to structural shocks are amplified, yielding smaller residual variations previously explained by the markup shocks. Consistent with this intuition, the contribution of price and wage markup shocks to output fluctuations is substantially smaller in the benchmark model than in the other two models, as will be shown in Section 3.3.3.

For the rest of the parameters, the estimation results are similar across three models. See Online Appendix B.3.

3.3 Business Cycle Implications

This section illustrates the business cycle implications of the normalized translog production function. Despite its parsimonious structure, our benchmark model features procyclical returns to scale and procyclical price markups that are consistent with the empirical findings in Section 2. Moreover, our benchmark model generates countercyclical labor shares and procyclical energy shares, which are consistent with the data, and maintains the countercyclical capital shares and procyclical profit shares, similar to the previous studies. Finally, we show that our framework leads to a considerable decrease in the importance of markup shocks on output fluctuations compared to the models with Cobb-Douglas production functions.

3.3.1 Returns to Scale and Price Markups

Armed with the estimated model using aggregate data, we revisit the returns to scale and price markup procyclicality results in Section 2. Given the normalized translog production function in Equation (3.1), which is denoted by $y_t(i) = f(k_t^s(i), l_t(i), e_t(i); l_t, e_t, \varepsilon_t^a)$, the returns to scale that firm i faces is given by $rts_t(i) \equiv \frac{\partial \log[f(\lambda k_t^s(i), \lambda l_t(i), \lambda e_t(i); l_t, e_t, \varepsilon_t^a)]}{\partial \log(\lambda)}|_{\lambda=1} = [1 + \beta_{el}(\hat{l}_t + \hat{e}_t)] \frac{y_t(i)+v}{y_t(i)}$. Note that

[18] Despite the differences in the Calvo parameters across the three models, they are all broadly consistent with the empirical estimates in Nakamura and Steinsson (2008) and Barattieri et al. (2014) and the model-based estimates in Justiniano et al. (2010, 2011).

[19] See Equations (A.88) and (A.92) in Online Appendix A for the price and wage Phillips curves.

firms take the aggregate labor l_t and energy e_t as given when they change their scales proportionally. Since the cross-sectional dispersion of $\log(y_t(i))$ is of the second order, the average returns to scale across firms, up to the first order, is given by

$$
\begin{aligned}
rts_t &= rts_t^{trans} \times rts_t^{fix} \\
&\equiv \left[1 + \beta_{el}\left(\hat{l}_t + \hat{e}_t\right)\right] \times \left(\frac{y_t + \upsilon}{y_t}\right),
\end{aligned}
\tag{3.10}
$$

where $rts_t^{trans} \equiv 1 + \beta_{el}\left(\hat{l}_t + \hat{e}_t\right)$ denotes the part of the returns to scale that arises from the translog structure, and $rts_t^{fix} \equiv \frac{y_t+\upsilon}{y_t}$ denotes the part of the returns to scale that emerges because of the presence of the fixed costs in production. Note that the rts_t is conceptually the same as our empirical measure of returns to scale rts_{jt} in Section 2.3 since both of them are based on the cross-sectional average of the firm-level returns to scale. As a result of the adjustment in the production function, compared to Equation (2.7), Equation (3.10) has only one production parameter β_{el} and additionally features the term relevant to the fixed costs in production rts_t^{fix}.

Clearly, the returns to scale varies over the business cycle with the translog structure and fixed costs. In the simplest case with Cobb-Douglas specification ($\beta_{el} = 0$) without the fixed costs ($\upsilon = 0$), the returns to scale is time-invariant: $rts_t = \beta_k + \beta_l + \beta_e = 1$. Allowing fixed costs in production recovers the returns to scale in Smets and Wouters (2007), $rts_t^{fix} = \frac{y_t+\upsilon}{y_t}$, which is countercyclical. The reason is because in a recession, firms face relatively higher fixed costs and thus larger returns to scale compared to the expansion. Additionally, allowing the translog structure makes the returns to scale more procyclical if inputs are complements ($\beta_{el} > 0$) and more countercyclical if inputs are substitutes ($\beta_{el} < 0$). Depending on the degree of input complementarity, the returns to scale can be procyclical with the presence of fixed costs.

For completeness, we define the aggregate returns to scale, which endogenizes the changes in aggregate labor and energy. Using the fact that y_t is obtained by aggregating $y_t(i)$ is the same as $f(k_t^s, l_t, e_t; l_t, e_t, \varepsilon_t^a)$ up to the first-order, the aggregate returns to scale is given by

$$
RTS_t = \left[1 + 2\beta_{el}\left(\hat{l}_t + \hat{e}_t\right)\right] \times \left(\frac{y_t + \upsilon}{y_t}\right),
\tag{3.11}
$$

where $RTS_t \equiv \frac{\partial \log[f(\lambda k_t^s, \lambda l_t, \lambda e_t; \lambda l_t, \lambda e_t, \varepsilon_t^a)]}{\partial \log(\lambda)}|_{\lambda=1}$. The aggregate returns to scale is derived under the assumption that all firms change their inputs proportionally such that the aggregate inputs change by the same proportion. The difference between RTS_t and $rts_t(i)$ is that for RTS_t, the exponents of the second-order term in Equation (3.1) endogenously vary with aggregate inputs when all firms adjust their inputs. In contrast, $rts_t(i)$ treats the exponents of the second-order term as constant because only firm i changes its inputs. This difference yields an additional term $\beta_{el}(\hat{l}_t + \hat{e}_t)$ to the aggregate returns to scale in Equation (3.11). Note that $RTS_t = rts_t$ without the translog structure

Table 4: Cyclicality of returns to scales and price markups

Correlation with $\log(y_t^{GDP})$	(1) $\log(rts_t^{trans})$	(2) $\log(rts_t^{fix})$	(3) $\log(rts_t)$	(4) $\log(RTS_t)$	(5) $\log(\Phi_t)$
HKL	0.74	-1	0.20	0.55	0.19
	(0.63, 0.85)	(-1, -1)	(-0.25, 0.54)	(0.30, 0.74)	(0.03, 0.26)
HKL-CD ($\beta_{el} = 0$)	-	-1	-1	-1	0.01
	-	(-1, -1)	(-1, -1)	(-1, -1)	(-0.14, 0.18)
Smets and Wouters (2007)	-	-1	-1	-1	0.00
	-	(-1, -1)	(-1, -1)	(-1, -1)	(-0.17, 0.18)

Notes: $rts_t^{trans} \equiv 1 + \beta_{el}(\hat{l}_t + \hat{e}_t)$ and $rts_t^{fix} \equiv \frac{y_t+v}{y_t}$ are the returns to scales that arise from the translog production function and fixed costs, respectively. $rts_t \equiv rts_t^{trans} \times rts_t^{fix}$ is the average returns to scale that each firm faces, $RTS_t \equiv [1 + 2\beta_{el}(\hat{l}_t + \hat{e}_t)]\frac{y_t+v}{y_t}$ is the aggregate returns to scale, and Φ_t is the gross price markup. All variables are log-linearized. For each of the three models, we report the correlations of each variable with y_t^{GDP} at the posterior mode and 90% credible intervals. HKL denotes our benchmark model with the translog production function, HKL-CD refers to the Cobb-Douglas specification with energy ($\beta_{el} = 0$), and Smets and Wouters (2007) is the Cobb-Douglas specification without energy ($\beta_{el} = 0, \beta_e = 0$).

($\beta_{el} = 0$).

Given the structure of our model, it is straightforward to recover the price markup. The changes in price markups can be most easily understood from the following expression of real marginal costs. By rearranging and combining the first-order conditions (3.2), (3.3), and (3.4) and the production function (3.1), we have:

$$\hat{\Phi}_t = -\widehat{mc}_t = \beta_{el}(\hat{l}_t + \hat{e}_t) - \beta_k \hat{r}_t^k - \beta_l \hat{w}_t - \beta_e \hat{p}_t^e + \varepsilon_t^a, \tag{3.12}$$

where Φ_t is the aggregate price markup, which depends on the translog part of the returns to scale, $\widehat{rts}_t^{trans} = \beta_{el}(\hat{l}_t + \hat{e}_t)$, real input prices $(\hat{r}_t^k, \hat{w}_t, \hat{p}_t^e)$, and productivity shocks (ε_t^a). The price markup rises when firms employ more complementary inputs, face a decrease in input prices, or experience positive productivity shocks. Equation (3.12) clarifies how traditional models link price markup cyclicality to price and wage rigidity. Previous studies have focused on the relative rigidities of prices and wages, which affect the cyclicality of $\hat{w}_t$. Consider the simple case of a Cobb-Douglas specification with only labor input; the price markup then becomes $\hat{\Phi}_t = -\beta_l \hat{w}_t + \varepsilon_t^a$. In this case, conditional on any shocks other than productivity shock, price markup cyclicality is governed by the cyclicality of real wage $\hat{w}_t$, which is tightly related to the relative rigidities of prices and wages. More generally, in previous studies, the cyclicality of real input prices affects price markup cyclicality. In our setup, we identify a new term that additionally changes the price markup, $\beta_{el}(\hat{l}_t + \hat{e}_t)$, by adopting a more general production function that allows a flexible input substitution pattern.

Columns (1)-(4) in Table 4 show that the returns to scale is procyclical in our benchmark model (HKL), which is consistent with the results in Section 2, but it is countercyclical in the other models.

Column (1) shows the cyclicality of the new returns to scale term that arises from the translog structure $rts_t^{trans} \equiv 1 + \beta_{el}(\hat{l}_t + \hat{e}_t)$. Given the posterior mode of $\beta_{el} = 0.042$, the correlation of rts_t^{trans} and GDP is 0.74; in boom periods, the larger use of each input generates synergies and makes the economy produce more. In contrast, as shown in column (2), rts_t^{fix} features perfect or nearly perfect countercyclical returns to scale in all three models as the fixed costs become relatively larger in recession.[20] Column (3) indicates that the total average returns to scale rts_t is mildly procyclical in our benchmark model, which shows that the procyclical effect of the translog structure dominates the countercyclical effect of the fixed costs. The aggregate returns to scale RTS_t features stronger procyclicality due to the endogenous movement in aggregate labor and energy, as appears in column (4). In the other models with Cobb-Douglas production functions, the returns to scale equals rts_t^{fix} and is countercyclical.

Column (5) shows that the price markup is procyclical in our benchmark model (HKL), as we hinted in Section 2. This procyclicality works through the new part of the returns to scale, $\widehat{rts}^{trans}$. In expansion, when firms utilize more labor and energy, the complementarity between these inputs leads to an increase in the marginal product of inputs, which in turn decreases the real marginal costs and increases the price markup. Without this mechanism, price markups are acyclical in our model, as in the models with Cobb-Douglas specification, because the net effect of real input price cyclicality on price markup cyclicality is negligible.

3.3.2 Factor and Profit Shares

Our translog production function has a novel implication on the income distribution across factors and profits through its effects on the first-order conditions. By rewriting Equations (3.2)-(3.4) in terms of factor shares, to the first-order approximation, we have:

$$s_t^e \equiv \frac{p_t^e e_t}{y_t^{GDP}} = \frac{1}{\Phi_t}\frac{y_t + \upsilon}{y_t^{GDP}}(\beta_e + \beta_{el}\hat{l}_t), \tag{3.13}$$

$$s_t^l \equiv \frac{w_t l_t}{y_t^{GDP}} = \frac{1}{\Phi_t}\frac{y_t + \upsilon}{y_t^{GDP}}(\beta_l + \beta_{el}\hat{e}_t), \tag{3.14}$$

$$s_t^k \equiv \frac{r_t k_t}{y_t^{GDP}} = \frac{1}{\Phi_t}\frac{y_t + \upsilon}{y_t^{GDP}}\beta_k, \tag{3.15}$$

[20] In the Smets and Wouters (2007) model, $\log(y_t^{GDP})$ and $\log\left(\frac{y_t+\upsilon}{y_t}\right)$ are perfectly negatively correlated because $y_t^{GDP} = y_t$. In models with energy input, although $y_t^{GDP} = y_t - (e_t - \phi_s e_t^s)$, we still observe the nearly perfect negative correlation between $\log(y_t^{GDP})$ and $\log\left(\frac{y_t+\upsilon}{y_t}\right)$. This is because the difference between y_t and y_t^{GDP}, the net energy import $e_t - \phi_e e_t^s$, is negligible relative to y_t.

Table 5: Cyclicality of factor and profit shares

Correlation with $\log(y_t^{GDP})$	(1) Energy Shares	(2) Labor Shares	(3) Capital Shares	(4) Profit Shares
HKL	0.46	-0.41	-0.47	0.42
	(0.26, 0.61)	(-0.50, -0.25)	(-0.55, -0.33)	(0.25, 0.50)
HKL-CD ($\beta_{el} = 0$)	-0.39	-0.39	-0.39	0.39
	(-0.48, -0.21)	(-0.48, -0.21)	(-0.48, -0.21)	(0.21, 0.48)
Smets and Wouters	-	-0.41	-0.41	0.41
	-	(-0.51, -0.22)	(-0.51, -0.22)	(0.22, 0.51)
Data	0.50	-0.38	-	-

Notes: For each of the three models, we report both the correlations of the logarithm of each variable with $\log(y_t^{GDP})$ at the posterior mode and the 90% credible intervals. We compute the empirical energy and labor shares based on the data used for the Bayesian estimation in Section 3.2 and employ the Baxter and King (1999) filter with a periodicity of cycles between 6 and 32 quarters. Following Gorodnichenko and Ng (2010), we apply the same filter to the model variables and calculate the correlation coefficients by using the representation in Croux et al. (2001, Equation (8)). HKL denotes our benchmark model with the translog production function, HKL-CD refers to the Cobb-Douglas specification with energy ($\beta_{el} = 0$), and Smets and Wouters (2007) is the Cobb-Douglas specification without energy ($\beta_{el} = 0, \beta_e = 0$).

where s_t^e, s_t^l, and s_t^k are energy, labor, and capital shares, respectively. Clearly, the translog structure and energy imports generate two changes in the factor share equations from those in Smets and Wouters (2007). First, the translog production function induces the extra terms in the energy and labor shares, $\beta_{el}\hat{l}_t$ and $\beta_{el}\hat{e}_t$, which constitute the cyclical components of rts^{trans}. Second, the imported energy used in production allows us to consider the energy shares and makes y_t^{GDP} marginally different from y_t by net energy imports. The smaller inverse input wedge—the multiplication of the price markup (Φ_t) and the inverse of the fixed-cost relevant term $(1/\left(\frac{y_t+\upsilon}{y_t^{GDP}}\right))$—and the larger complementarity ($\beta_{el} > 0$) leads to a greater factor demand and factor shares in expansions.

By combining the input share expressions (3.13)-(3.15) with the definition of the returns to scale (3.10), the profit share becomes:

$$s_t^{\Pi} = \left(1 - \frac{rts_t}{\Phi_t}\right)\frac{y_t}{y_t^{GDP}}, \tag{3.16}$$

where $s_t^{\Pi} \equiv \frac{\Pi_t}{y_t^{GDP}}$ with the aggregate profit $\Pi_t \equiv y_t - w_t l_t - r_t^k k_t^s - p_t^e e_t$. Thus, our model generates a relationship among the profit share, price markup, and returns to scale, consistent with the previous studies of the returns to scale and the empirical expression (2.8). Holding everything else constant, larger price markup or smaller returns to scale leads to smaller total input shares and larger profit shares. Note that the returns to scale cyclicality can substitute for the price markup cyclicality in explaining the same profit share dynamics.

Our benchmark model better matches the energy input share in the data than the model with the Cobb-Douglas production function. As shown in column (1) of Table 5, our benchmark model features procyclical energy shares with a correlation of 0.46, which is both qualitatively and quantitatively consistent with the data.[21] The procyclical energy shares arise from the complementarity between labor and energy ($\beta_{el} > 0$), where a positive β_{el} is motivated by the empirical analyses (Section 2) and is estimated by the Bayesian method (Section 3.2). This procyclical effect of complementarity dominates the countercyclical effects of the inverse price markup ($1/\Phi_t$) and the fixed costs ($\frac{y_t+v}{y_t^{GDP}}$) on energy shares. In contrast, given the Cobb-Douglas production function, the energy share is identically countercyclical to the other factor shares. Its simple assumption on the functional form does not allow us to generate different input share cyclicality as in the data.

At the same time, our benchmark model maintains countercyclical labor and capital shares, which are consistent with the models with Cobb-Douglas specifications. As presented in columns (2) and (3), the cyclicality of labor and capital shares is similar across the three different models. In particular, the labor share cyclicality is comparable to the data and to what are reported in Karabarbounis (2014, Table 7) and Rios-Rull and Santaeulalia-Llopis (2010, Table 2). Both the labor and capital shares are countercyclical in all models because of the fixed costs in production, which generate a countercyclical wedge and depress the labor and capital demand in expansions. For the labor share, although our benchmark model has an additional procyclical complementarity term ($\beta_l + \beta_{el}\hat{e}_t$), it is canceled out by the procyclical price markup effect (Φ_t), which results in countercyclical labor shares.[22] The capital share is slightly more countercyclical in our benchmark model relative to the capital share in the two other models due to procyclical price markups.

Finally, the profit share is procyclical in all three models, as shown in column (4), but the underlying mechanism is largely different depending on the production function specification. The profit share procyclicality in our benchmark model arises from the procyclical price markups. Since firms generate more price markups in expansion as a result of the depressed marginal costs with input complementarity, their profit shares rise at the same time. The returns to scale is procyclical in this model and mitigates the profit procyclicality, but this effect is not strong enough to make profit shares countercyclical. In contrast, in the models with a Cobb-Douglas production function, the profit share procyclicality relies on the countercyclical returns to scale. Without the complementarity, in expansion, the fixed costs depress the returns to scale, which in turn decrease all input shares and increase the profit shares. The price markup is acyclical in these models and does not contribute to the procyclical profit. Note that the difference between y_t and y_t^{GDP} is negligible and does not contribute to the cyclicality of profit shares in any model.

[21] For $US\ ENERGY_t$, we use monthly US industrial energy usage data from U.S. Energy Information Administration (2021). We seasonally adjust this series by using X-13 ARIMA-SEATS and aggregate it to a quarterly measure. Because this monthly measure is available from 1973, our sample for Table 5 spans from 1973:q1 to 2004:q4.

[22] Note that the magnitude of procyclical variation in the complementarity term in the labor share ($\beta_l + \beta_{el}\hat{e}_t$) is smaller than that in the energy share, which results in a procyclical energy share but a countercyclical labor share. This is because the energy share β_e is small, and with log-linearization, $\beta_e + \beta_{el}\hat{l}_t$ becomes $\frac{\beta_{el}}{\beta_e}\hat{l}_t$, which generates a large procyclical component in the energy share.

In addition to assessing the unconditional moments, we further verify our benchmark model by assessing the conditional moments in Online Appendix B.4. Given that our theoretical mechanism is centered around the complementarity between labor and energy, we compare the model responses of labor and energy to monetary and fiscal policy shocks with the corresponding empirical impulse responses, which are estimated by using the identified structural shocks in Romer and Romer (2004) and Auerbach and Gorodnichenko (2012). We find that the model responses are largely consistent with the empirical responses in our benchmark model. Note that this consistency holds without directly matching the dynamics of US industrial energy usage e_t in the Bayesian estimation in Section 3.2.

3.3.3 Contribution of Markup Shocks

After observing the changes in the returns to scale, price markup, and factor share cyclicality, which are induced by the translog production function, this section seeks to understand the relative importance of different driving forces of business cycles. For this purpose, we compute forecast variance decompositions (FEVDs) and impulse responses of the three models with different production functions.

The translog specification substantially decreases the contribution of price and wage markup shocks to business cycles. Table 6 presents the FEVDs of output and labor at ten quarter horizon based on all three models. While introducing the energy input into a Cobb-Douglas technology (HKL-CD) does not meaningfully change the FEVD from that of the Smets and Wouters (2007) model, additionally introducing the translog specification (HKL) alters the results substantially. The most notable change is the contribution of markup shocks. Price and wage markup shocks explain 28% and 26% of output fluctuations in the Smets and Wouters (2007) model and the model without complementarity (HKL-CD), respectively. In contrast, the corresponding FEVD decreases to 15% in our benchmark model. The results for labor are similar. The FEVD of labor with respect to markup shocks slightly decreases from 45% to 44% when we introduce an energy input, but it becomes 31% when we incorporate the input complementarity between labor and energy. At alternative horizons, we still observe that markup shocks are less important determinants of output and labor in our benchmark model than in the two other models. As a result, the other structural shocks, such as shocks to productivity, demand, and energy, become more important drivers of business cycles in our benchmark model than in the other models. See Online Appendix B.5 for additional results.

Given the substantial changes in the contribution of markup shocks, Figure 2 presents the impulse responses of output to price and wage markup shocks of all three models to investigate the underlying mechanism behind the FEVDs. The top (bottom) panels show the responses to a one-standard-deviation contractionary price (wage) markup shock, and the solid, dash-dotted, and dashed lines represent the results based on our benchmark model (HKL), the model without input

Table 6: Forecast error variance decomposition of output and labor (10 quarters)

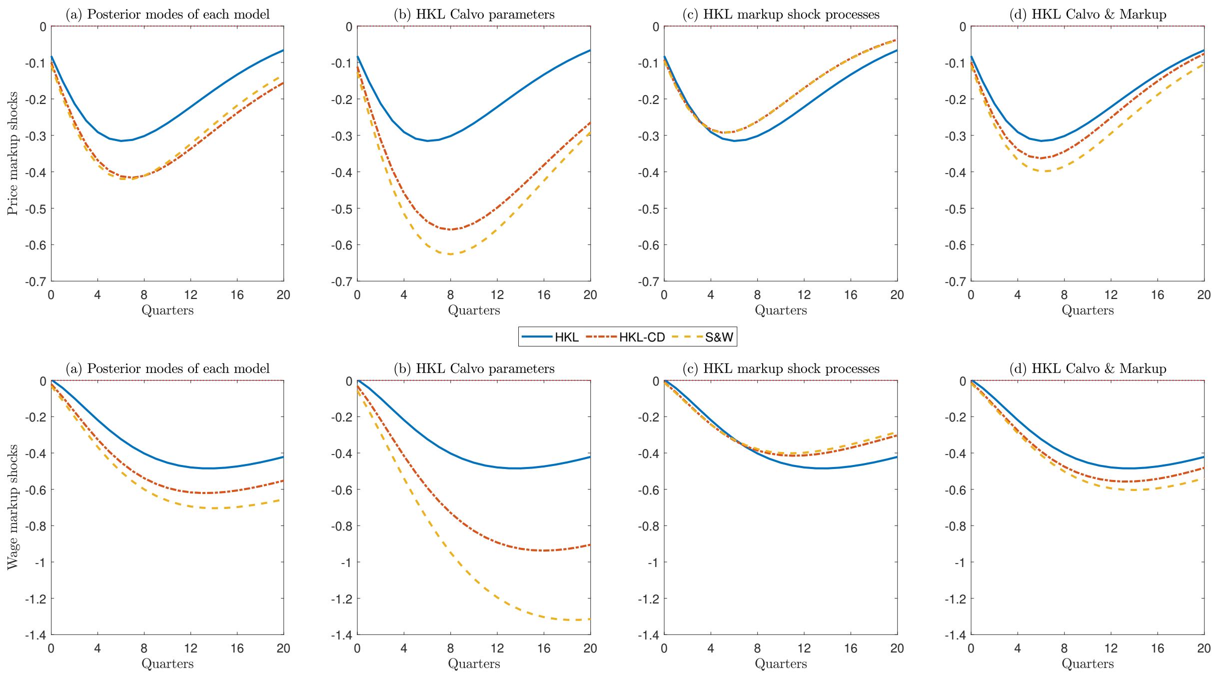
	Output ($\log y_t^{GDP}$)			Labor ($\log l_t$)		
	HKL	HKL-CD	S&W	HKL	HKL-CD	S&W
Panel A						
Productivity (neutral)	0.36	0.34	0.32	0.05	0.04	0.04
Risk premium	0.05	0.04	0.04	0.09	0.07	0.06
Government spending	0.09	0.08	0.08	0.17	0.15	0.15
Investment-specific productivity	0.22	0.21	0.22	0.24	0.21	0.21
Monetary policy	0.08	0.06	0.06	0.13	0.09	0.09
Price markup	0.07	0.11	0.11	0.08	0.13	0.12
Wage markup	0.08	0.15	0.17	0.22	0.31	0.33
Energy demand	0.02	0.01	-	0.01	0.00	-
Energy supply	0.03	0.01	-	0.01	0.00	-
Panel B						
Productivity shocks	0.58	0.54	0.54	0.29	0.25	0.25
Demand shocks	0.21	0.19	0.19	0.39	0.31	0.30
Markup shocks	0.15	0.26	0.28	0.31	0.44	0.45
Energy shocks	0.06	0.01	-	0.01	0.01	-

Notes: Panel A decomposes the forecast error variances into the contributions of the 9 structural shocks in the model. Panel B summarizes the FEVDs of different types of shocks. The productivity shocks include the neutral and investment-specific productivity shocks, the demand shocks include the risk premium, government spending, and monetary policy shocks, the markups shocks include the price and wage markup shocks, and the energy shocks include the energy demand and supply shocks. HKL denotes our benchmark model with the translog production function, HKL-CD refers to the Cobb-Douglas specification with energy ($\beta_{el} = 0$), and Smets and Wouters (2007) features the Cobb-Douglas production function without energy ($\beta_{el} = 0, \beta_e = 0$).

complementarity (HKL-CD), and the Smets and Wouters (2007) model (S&W), respectively. Column (a) presents the impulse responses at the posterior mode of the three models. Columns (b)-(d) set different subsets of parameters at the posterior mode of HKL to investigate why the price markup shocks have smaller effects in our benchmark model. We fix two categories of parameters reported in Table 3, which have noticeably different posterior modes across the models: (i) Calvo price and wage stickiness parameters and (ii) persistence of markup shocks. For all three models, column (b) uses the Calvo parameters, column (c) uses the markup shock parameters, and column (d) uses both the Calvo parameters and markup shock parameters in our benchmark model. In Online Appendix B.6, we separately test whether the positive input complementarity parameter ($\beta_{el} > 0$) can depress the roles of markup shocks but do not find the supporting evidence. It mostly substitutes the countercyclical variations in price markups in the previous models, as discussed with the profit share cyclicality result in Section 3.3.2.

Consistent with Table 6, Figure 2 column (a) clearly shows that the markup shocks have

Figure 2: Impulse responses of output to price and wage markup shocks



Notes. The two panels in column (a) show the responses of output to a one-standard-deviation contractionary price and wage markup shock, respectively. The solid, dashdotted, and dashed lines represent the results based on our benchmark model (HKL), the model without input complementarity (HKL-CD), and the Smets and Wouters (2007) model (S&W), respectively. We use the posterior modes of each model. For column (b), we replace the Calvo sticky price and wage parameters of HKL-CD and S&W to the posterior mode of the HKL model. Column (c) is based on the estimated price and wage markup shock processes of HKL in Table 3. Finally, column (d) uses HKL posterior modes for both Calvo parameters and markup shock processes.

substantially smaller effects on output in our benchmark model than that in the two other models. The peak effects of a one-standard-deviation price markup shock on output are 0.32% (HKL), 0.42% (HKL-CD) and 0.42% (S&W) in absolute value. Similarly, for wage markup shocks, the peak effects are 0.48% (HKL), 0.62% (HKL-CD) and 0.70% (S&W). The smaller effects of both price and wage markup shocks with the translog production function are similar for investment, consumption, and labor, as shown in Appendix D.

By comparing Figure 2 columns (b)-(d), we find that less persistent price markup shocks are mostly responsible for the depressed role of markup shocks in our benchmark model. As shown in column (b), fixing the price and wage rigidity across the three models makes the impulse response generated from HKL-CD and Smets and Wouters (2007) to deviate even more from that of our benchmark model, highlighting that the larger rigidity in our benchmark model does not decrease the importance of the markup shocks. However, the impulse responses of output are nearly identical across the three models when we use the same markup shock parameters, as shown in column (c). These results show that the decrease in markup shock parameters (i.e., lower persistence of markup shock processes) is essential for generating the smaller contribution of markup shocks. Column (d) fixes both Calvo and markup shock parameters, and the impulse responses are analogous to those in column (a).

Allowing the procyclical returns to scale leads to the less persistent markup shocks through the flatter price and wage Philips curves. As shown in Table 4, our benchmark model features the procyclical price markup or the countercyclical real marginal costs because of the input complementarity. The large countercyclical fluctuations in the marginal costs per se do not amplify the dynamics of aggregate variables. However, it makes our Bayesian estimator select stickier prices and less responsive inflation to the real marginal costs to match the empirical correlation of inflation and GDP. Correspondingly, nominal wages become stickier to match the real wage cyclicality in the data. The larger price and wage rigidity, in turn, makes real variables respond more to structural shocks, as shown in Figure 2 column (b), and absorb previously unexplained variations that were attributed to price and markup shocks.

4 Conclusion

This paper studies business cycles with a translog production function. Our empirical analyses suggest that there is complementarity between labor and energy, which leads to procyclical returns to scale. Our empirical evidence is not compatible with the tightly parameterized production functions commonly used in the literature. Thus, we introduce the normalized translog production function into a standard medium-scale DSGE model and re-estimate the substitution parameters within the structure of our model. Our model rationalizes procyclical returns to scale, procyclical price markups, countercyclical labor shares, procyclical energy shares, and procyclical profit shares. Furthermore, we document that the contribution of price and wage markup shocks to output fluctuations in our model is substantially smaller than that in the standard Smets and Wouters (2007) model. The complementarity between labor and energy and the corresponding procyclical returns to scale are central to the theoretical mechanism behind the results.

Our work underscores the need to employ more general forms of production functions in business cycle research. Further efforts to utilize a general functional form will extend the understanding of business cycles.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Structural Identification of Production Functions," *Econometrica*, 2015, *83* (6), 2411–2451.

Anderson, Eric, Sergio Rebelo, and Arlene Wong, "Markups Across Space and Time," *Working paper*, 2020.

Antras, Pol, "Is the U.S. aggregate production function Cobb-Douglas? New estimates of the elasticity of substitution," *Contributions to Macroeconomics*, 2004, *4* (1).

Arellano, Cristina, Yan Bai, and Patrick J. Kehoe, "Financial Frictions and Fluctuations in Volatility," *Journal of Political Economy*, 2019, *127*, 2049–2103.

Arellano, Manuel and Stephen Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies*, 04 1991, *58* (2), 277–297.

Atalay, Enghin, "How Important Are Sectoral Shocks," *American Economic Journal: Macroeconomics*, 2017, *9* (4), 254–280.

Auerbach, Alan J. and Yuriy Gorodnichenko, "Measuring the output responses to fiscal policy," *American Economic Journal: Economic Policy*, 2012, *4* (2), 1–27.

Backus, David K. and Mario J. Crucini, "Oil prices and the terms of trade," *Journal of International Economics*, 2000, *50* (1), 185–213.

Balke, Nathan S. and Stephen P.A. Brown, "Oil supply shocks and the US economy: An estimated DSGE model," *Energy Policy*, 2018, *116*, 357–372.

Baqaee, David Rezza and Emmanuel Farhi, "The Darwinian Returns to Scale," *Working paper*, 2021.

Barattieri, Alessandro, Susanto Basu, and Peter Gottschalk, "Some evidence on the importance of sticky wages," *American Economic Journal: Macroeconomics*, 2014, *6* (1), 70–101.

Bartelsman, Eric J., Randy A. Becker, and Wayne Gray, "The NBER-CES Manufacturing Industry Database," 2000. National Bureau of Economic Research Working Paper, Cambridge, MA.

Basak, Suleyman and Anna Pavlova, "A model of financialization of commodities," *Journal of Finance*, 2016, *71* (4), 1511–1556.

Basu, Susanto, "Intermediate goods and business cycles," *American Economic Review*, 1995, *85* (3), 512–531.

_ and John Fernald, "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, 1997, *105* (2), 249–283.

Baxter, Marianne and Robert G. King, "Productive externalities and business cycles," *Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis, Discussion Paper 53*, 1991.

_ and _ , "Measuring business cycles: approximate band-pass filters for economic time series," *Review of Economics and Statistics*, 1999, *81* (4), 575–593.

Benhabib, Jess and Roger E. A. Farmer, "Indeterminancy and Increasing Returns," *Journal of Economic Theory*, 1994, *63*, 19–41.

_ **and** _ , "Indeterminancy and Sector-Specific Externalities," *Journal of Monetary Economics*, 1996, *37*, 421–443.

Berger, David, Kyle Herkenhoff, and Simon Mongey, "Labor Market Power," 2019. Working Paper.

Berman, Nicolas, Antoine Berthou, and Jérôme Héricourt, "Export Dynamics and Sales at Home," *Journal of International Economics*, 2015, *96* (2), 298–310.

Bigio, Saki and Jennifer La'O, "Distortions in Production Networks," 2020. Working Paper.

Bils, Mark, "The Cyclical Behavior of Marginal Cost and Price," *American Economic Review*, 1987, *77* (5), 838–855.

_ **, Peter J. Klenow, and Benjamin A. Malin**, "Testing for Keynesian Labor Demand," in "NBER Macroeconomics Annual 2013," University of Chicago Press, 2013, pp. 311–349.

_ **,** _ **, and** _ , "Resurrecting the Role of the Product Market Wedge in Recessions," *American Economic Review*, 2018, *108* (4-5), 1118–1146.

Blundell, Richard and Stephen Bond, "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 1998, *87* (1), 115–143.

BP Energy, "BP Statistical Review of World Energy," 2020. June.

Cantore, Cristiano and Paul Levine, "Getting normalization right: Dealing with 'dimensional constants' in macroeconomics," *Journal of Economic Dynamics and Control*, 2012, *36* (12), 1931–1949.

_ **, Miguel León-Ledesma, Peter McAdam, and Alpo Willman**, "Shocking Stuff: Technology, Hours, and Factor Substitution," *Journal of the European Economic Association*, 2014, *12* (1), 108–128.

_ **, Paul Levine, Joseph Pearlman, and Bo Yang**, "CES technology and business cycle fluctuations," *Journal of Economic Dynamics and Control*, 2015, *61*, 133–151.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan, "Business Cycle Accounting," *Econometrica*, 2007, *75* (3), 781–836.

_ **,** _ **, and** _ , "New Keynesian Models: Not Yet Useful for Policy Analysis," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 242–266.

Chrinko, Robert S., "σ:The Long and Short of It," *Journal of Macroeconomics*, 2008, *30* (2), 671–686.

Christensen, Lauritis, Dale Jorgensen, and Lawrence Lau, "Transcendental logarithmic Production Frontier," *Review of Economics and Statistics*, 1973, *55* (1), 28–45.

_ **,** _ **, and** _ , "Transcendental logarithmic utility functions," *American Economic Review*, 1975, *65* (3), 367–383.

Collard-Wexler, Allan and Jan De Loecker, "Production Function Estimation and Capital Measurement Error," *Working paper*, 2020.

Croux, Christophe, Mario Forni, and Lucrezia Reichlin, "A measure of comovement for economic variables: Theory and empirics," *Review of Economics and Statistics*, 2001, *83* (2), 232–241.

De Jong, Frits J., *Dimensional Analysis for Economists*, North Holland, 1967.

De Loecker, Jan and Pinelopi K. Goldberg, "Firm Performance in a Global Market," *Annual Review of Economics*, 2014, *6*, 201–227.

_ , Jan Eeckhout, and Gabriel Unger, "The Rise of Market Power and the Macroeconomic Implications," *Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

_ , Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik, "Prices, Markups and Trade Reform," *Econometrica*, 2016, *84* (2), 445–510.

Dhawan, Rajeev and Karsten Jeske, "Energy price shocks and the macroeconomy: the role of consumer durables," *Journal of Money, Credit and Banking*, 2008, *40* (7), 1357–1377.

Doraszelski, Ulrich and Jordi Jaumandreu, "R&D and Productivity: Estimating Endogenous Productivity," *Review of Economic Studies*, 2013, *126* (3), 1027–1084.

Epifani, Paolo and Gino Gancia, "Increasing returns, imperfect competition, and factor prices," *Review of Economics and Statistics*, 2006, *88* (4), 583–598.

Evans, Charles L., "Productivity shocks and real business cycles," *Journal of Monetary Economics*, 1992, *29* (2), 191–208.

Fernald, John, "A quarterly, utilization-adjusted series on total factor productivity," 2014. Federal Reserve Bank of San Francisco Working Paper 2012-19.

Finn, Mary G., "Perfect competition and the effects of energy price increases on economic activity," *Journal of Money, Credit and Banking*, 2000, *32* (3), 400–416.

Galí, Jordi, *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*, Princeton University Press, 2015.

Gali, Jordi, Mark Gertler, and J. David Lopez-Salido, "Markups, Gaps, and The Welfare Costs of Business Fluctuations," *Review of Economics and Statistics*, 2007, *89* (1), 44–59.

Gandhi, Amit, Salvador Navarro, and David Rivers, "On the Identification of Production Functions," *Journal of Political Economy*, 2020, *128* (8), 2973–3016.

Gilchrist, Simon, Raphael Schoenle, Jae W. Sim, and Egon Zakrajsek, "Inflation Dynamics During the Financial Crisis," *American Economic Review*, 2017, *107* (3), 785–823.

Goodfriend, Marvin and Robert King, "The New Neoclassical Synthesis and the Role of Monetary Policy," in "NBER Macroeconomics Annual 1997, Volume 12," MIT Press, 1997, pp. 231–296.

Gorodnichenko, Yuriy and Serena Ng, "Estimation of DSGE models when the data are persistent," *Journal of Monetary Economics*, 2010, *57* (3), 325–340.

Gutierrez, German and Thomas Philippon, "Declining Competition and Investment in the U.S.," 2017. NBER Working Paper No. 23583.

Hall, Robert E., "The relation between price and marginal cost in US industry," *Journal of Political*

Economy, 1988, *96* (5), 921–947.

_ , "Invariance properties of Solow's Productivity Residual," in Peter Diamond, ed., *Growth/Productivity/Unemployment: Essays to Celebrate Bob Solow's Birthday*, Cambridge, MA: MIT Press, 1990, pp. 72–112.

_ , "Measuring Factor Adjustment Costs," *Quarterly Journal of Economics*, 2004, *119* (3), 899–927.

_ , "What the cyclical response of advertising reveals about markups and other macroeconomic wedges," Working Paper 18370, National Bureau of Economic Research, Inc. October 2013.

Hamilton, James D., "Understanding crude oil prices," *Energy Journal*, 2009, *30* (2), 179–206.

Hanoch, Giora, "The elasticity of scale and the shape of average costs," *American Economic Review*, 1975, *65* (3), 492–497.

Hansen, Lars Peter, "Large sample properties of generalized method of moments estimators," *Econometrica*, 1982, *50* (4), 1029–1054.

Hassler, John, Per Krusell, and Conny Olovsson, "Directed technical change as a response to natural-resource scarcity," *Working paper*, 2019.

Hottman, Colin J., Stephen J. Redding, and David E. Weinstein, "Quantifying the Sources of Firm Heterogeneity," *Quarterly Journal of Economics*, 2016, *131* (3), 1291–1364.

Jermann, Urban and Vincenzo Quadrini, "Macroeconomic Effects of Financial Shocks," *American Economic Review*, 2012, *102* (1), 238–271.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti, "Investment shocks and business cycles," *Journal of Monetary Economics*, 2010, *57* (2), 132–145.

_ , _ , and _ , "Investment shocks and the relative price of investment," *Review of Economic Dynamics*, 2011, *14* (1), 102–121.

Kaldor, Nicholas, "A Model of Economic Growth," *Economic Journal*, 1957, *67* (268), 591–624.

Karabarbounis, Loukas, "The labor wedge: MRS vs. MPN," *Review of Economic Dynamics*, 2014, *17* (2), 206–223.

_ **and Brent Neiman**, "The Global Decline of the Labor Share," *Quarterly Journal of Economics*, 2014, *129* (1), 61–103.

Kilian, Lutz, "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market," *American Economic Review*, 2009, *99* (3), 1053–1069.

_ , "Oil Price Shocks: Causes and Consequences," *Annual Review of Resource Economics*, 2014, *6* (1), 133–154.

_ **and Daniel P. Murphy**, "Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VAR models," *Journal of the European Economic Association*, 2012, *10* (5), 1166–1188.

Koh, Dongya and Raül Santaeulàlia-Llopis, "Countercyclical Elasticity of Substitution," 2017. Working Paper.

Levinsohn, James A. and Amil Petrin, "Estimating Production Functions Using Inputs to

Control for Unobservables," *Review of Economic Studies*, 2003, *70* (2), 317–340.

Nakamura, Emi and Jon Steinsson, "Five Facts about Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics*, 2008, *123* (4), 1415–1464.

_ **and** _ , "Monetary Non-Neutrality in a Multi-Sector Menu Cost Model," *Quarterly Journal of Economics*, 2010, *125* (3), 961–1013.

Nekarda, Christopher J. and Valerie A. Ramey, "The cyclical behavior of the price-cost markup," *Journal of Money, Credit and Banking*, 2020, *52* (S2), 319–353.

Oberfield, Ezra and Devesh Raval, "Micro data and macro technology," *Econometrica*, 2021, *89* (2), 703–732.

Olley, Steve G. and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Rajan, Raghuram G. and Luigi Zingales, "Financial Dependence and Growth," *American Economic Review*, 1998, *88* (3), 559–586.

Raval, Devesh R., "The micro elasticity of substitution and non-neutral technology," *RAND Journal of Economics*, 2019, *50* (1), 147–167.

Ravn, Morten O., Stephanie Schmitt-Grohé, and Martin Uribe, "Deep Habits," *Review of Economic Studies*, 2006, *73*, 195–218.

Rios-Rull, Jose-Victor and Raul Santaeulalia-Llopis, "Redistributive shocks and productivity shocks," *Journal of Monetary Economics*, 2010, *57* (8), 931–948.

Romer, Christina D. and David H. Romer, "A new measure of monetary shocks: Derivation and implications," *American Economic Review*, 2004, *94* (4), 1055–1084.

Rotemberg, Julio J., "Comment on Testing for Keynesian Labor Demand," in "NBER Macroeconomics Annual 2013," Vol. 27 (1), University of Chicago Press, 2013, pp. 362–370.

_ **and Michael Woodford**, "Markups and the Business Cycle," in "NBER Macroeconomics Annual 1991," MIT Press, 1991, pp. 63–129.

_ **and** _ , "Imperfect competition and the effects of energy price increases on economic activity," *Journal of Money, Credit, and Banking*, 1996, *28* (4), 550–577.

_ **and** _ , "The Cyclical Behavior of Prices and Costs," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, North Holland: Elsevier Science, 1999, chapter 16, pp. 1051–1135.

Schmitt-Grohé, Stephanie, "Endogenous Business Cycles and the Dyanmics of Output, Hours, and Consumption," *American Economic Review*, 2000, *90* (5), 1136–1159.

Sims, Christopher A., "Solving linear rational expectations models," *Computational Economics*, 2002, *20* (1-2), 1–20.

Smets, Frank and Rafael Wouters, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 2007, *97* (3), 586–606.

Stroebel, Johannes and Joseph Vavra, "House Prices, Local Demand, and Retail Prices," *Journal of Political Economy*, 2019, *127* (3), 1391–1436.

Syverson, Chad, "What determines productivity?," *Journal of Economic Literature*, 2011, *49* (2), 326–365.

U.S. Energy Information Administration, *Monthly Energy Review February 2021*, Washington, DC: Office of Energy Statistics, U.S. Department of Energy, 2021.

Appendix A NBER-CES: Summary Statistics

Table A.1 reports the basic summary statistics for the main dataset we use in the paper (NBER-CES manufacturing database). The data covers 473 6-digit NAICS manufacturing industry in 1958-2009. The quantity index of five inputs are measured with the energy (*energy/pien*), material excluding energy (*(matcost/pimat) - (energy/pien)*, since the units in the value and price of material and energy are identical), production workers (*prode*), non-production workers (*emp - prode*), and capital (*cap*). The energy share is measured as the share of the value of shipment (*energy/vship*). We use the input prices for the instrumental variables, and they are the energy price deflator (*pien*), material price deflator (*pimat*), production worker wage (*prodw/prode*), worker wage (*pay/emp*), and investment deflator (*piinv*). Note that we do not construct the price indexes separately for the material excluding the energy input and the non-production worker, as they are not given in the original dataset and requires more subtraction and addition, which would like to exacerbate the measurement error problems in such variables. For example, we find that if we measure the non-production worker wage by ourselves (*(pay-prodw)/(emp-prode)*), we observe that this variable has large negative values. Table A.2 reports the first-stage regression results for the instrumental variables used in this paper.

Table A.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
NAICS 6-digit Codes	327009.662	8889.717	311111	339999	24596
Year ranges from 58 to 09	1983.5	15.009	1958	2009	24596
Total employment in 1000s	34.814	45.053	0.2	559.9	24167
Total payroll in $1m	735.896	1252.867	2.9	16162.9	24167
Production workers in 1000s	25.423	33.651	0.2	459.9	24167
Production worker hours in 1m	50.64	66.823	0.3	904.1	24167
Production worker wages in $1m	443.796	736.828	1.8	10475.2	24167
Total value of shipments in $1m	4799.495	13196.309	19.3	732728.4	24167
Total cost of materials in $1m	2620.97	9721.219	8.800	648048.4	24167
Total value added in $1m	2190.409	4710.187	9.700	111187.9	24167
Total capital expendture in $1m	156.655	462.108	0.1	17601.6	24167
End-of-year inventories in $1m	585.429	1433.618	1.3	40084.9	24162
Cost of electric & fuels in $1m	97.69	360.53	0.1	14201.5	24167
Total real capital stock in $1m	2757.95	6388.03	4.1	120110.3	24167
Real capital: equipment in $1m	1664.517	4145.339	1.9	88454.600	24167
Real capital: structures in $1m	1093.433	2418.355	2.2	38874.4	24167
Deflator for VSHIP 1997=1.000	0.799	1.552	0.044	47.409	24167
Deflator for MATCOST 1997=1.000	0.718	0.357	0.127	2.777	24167
Deflator for INVEST 1997=1.000	0.694	0.309	0.183	1.581	24167
Deflator for ENERGY 1997=1.000	0.695	0.402	0.087	2.233	24167
5-factor TFP annual growth rate	0.003	0.066	-0.642	1.387	23694
5-factor TFP index 1997=1.000	0.937	0.257	0.012	13.192	24167
4-factor TFP annual growth rate	0.003	0.066	-0.642	1.387	23694
4-factor TFP index 1997=1.000	0.936	0.257	0.011	13.193	24167
production worker hours per worker in 1000s	2.004	0.119	1.28	3.228	24167

Table A.2: Regressing Inputs on Input Prices

	(1) Energy	(2) Material	(3) Capital	(4) Worker^p	(5) Worker^{np}
energy price$_{(t-1)}$	-0.523***	0.158**	-0.166***	-0.054	-0.126**
	(0.091)	(0.071)	(0.049)	(0.064)	(0.058)
energy price$_{(t-2)}$	-0.542***	-0.202***	-0.224***	-0.292***	-0.240***
	(0.090)	(0.068)	(0.048)	(0.062)	(0.056)
material price$_{(t-1)}$	0.156**	-0.696***	-0.338***	0.085	-0.026
	(0.071)	(0.071)	(0.048)	(0.057)	(0.059)
material price$_{(t-2)}$	-0.183***	-0.033	0.015	0.081	0.111*
	(0.069)	(0.068)	(0.047)	(0.055)	(0.058)
investment price$_{(t-1)}$	-1.899***	-2.003***	-1.573***	-0.621**	-2.130***
	(0.313)	(0.349)	(0.216)	(0.286)	(0.294)
investment price$_{(t-2)}$	2.327***	1.601***	1.263***	1.094***	2.263***
	(0.317)	(0.350)	(0.216)	(0.287)	(0.296)
production worker wage$_{(t-1)}$	0.577***	0.380***	0.117	0.378***	0.047
	(0.134)	(0.135)	(0.078)	(0.123)	(0.107)
production worker wage$_{(t-2)}$	0.308**	0.028	0.067	0.104	-0.443***
	(0.141)	(0.139)	(0.080)	(0.125)	(0.109)
payroll per worker$_{(t-1)}$	-0.069	0.445***	0.265***	-0.529***	0.219*
	(0.158)	(0.158)	(0.091)	(0.149)	(0.129)
payroll per worker$_{(t-2)}$	-0.059	0.114	0.424***	-0.193	0.422***
	(0.166)	(0.161)	(0.092)	(0.148)	(0.128)
Constant	0.000	0.000	0.000	-0.000	-0.000
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
F-stat	135.29	135.12	159.30	57.06	27.53
Observations	23221	23221	23221	23221	23220

Note. Each double-demeaned input is regressed on all 10 double-demeaned lagged input prices used in Equation (2.6). "Energy" (*energy/pien*) is the quantity of energy, "Material" (*(matcost/pimat) - (energy/pien)*) is the quantity index of material that excludes the energy, "Capital" (*cap*) is the real capital, "Worker^p" (*prode*) is the number of production workers, and "Worker^{np}" (*emp - prode*) is the number of non-production workers. The "energy price" is the energy deflator (*pien*), the "material price" is the material deflator (*pimat*), the "investment price" is the investment deflator (*piinv*), the "production worker wage" is the production worker wage (*prodw/prode)*, and the "payroll per worker" is the average wage of all workers (*pay/emp*).

Appendix B Empirical Results: Robustness

B.1 Using Two-year and Three-year Lagged Input Prices

As a robustness exercise, we use two- and three-year lagged input prices as instrumental variables instead of one-year and two-year lagged input prices. Table B.1 presents the results. The key complementarity parameter between energy and production workers is still positive and statistically significant.

Table B.1: Estimation of Equation (2.6)

	IV (Lagged Input Prices)			OLS		
	(1)	(2)	(3)	(4)	(5)	(6)
energy	-0.286	-0.330**	-0.351**	0.843***	0.838***	0.816***
	(0.188)	(0.163)	(0.165)	(0.028)	(0.029)	(0.031)
production worker	1.255***	1.341***	1.272***	-0.326***	-0.323***	-0.464***
	(0.356)	(0.310)	(0.310)	(0.029)	(0.030)	(0.032)
non-production worker	-0.472*	-0.510**	-0.425*	-0.119***	-0.125***	-0.219***
	(0.246)	(0.244)	(0.221)	(0.022)	(0.022)	(0.025)
material	0.403*	0.386**		-0.409***	-0.420***	
	(0.235)	(0.175)		(0.025)	(0.026)	
capital	-0.070		0.449*	-0.032		-0.247***
	(0.403)		(0.239)	(0.021)		(0.026)
Observations	22747	22747	22747	22747	22747	22747

Note. The regression specifications are identical to those in Table 1. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

B.2 Allowing Additional Control Variables

This section explicitly considers the potential omitted variables that might be correlated with our instrumental variables in estimating Equation (2.6). For our baseline approach, we assume that the instrumental variables—the double-demeaned lagged input prices—are not correlated with the double-demeaned wedge term τ. This assumption is consistent in a model that does not have an industry-time-varying energy wedge and the non-parametric estimation strategy employed in Gandhi et al. (2020). However, in a more general class of models, our assumption could be violated. For example, in a dynamic model that allows the heterogeneity of output price rigidity and variation in input prices across sectors, the industry-time-varying lagged input prices could be correlated with the industry-time-varying contemporaneous price-cost markups, which is a part of the industry-time-varying contemporaneous wedge τ^e.

Given such a concern, we control the potential omitted variables directly and re-estimate the first-order conditions. First, we measure the industry-level market power and directly control it to address such potential omitted variables. We carefully follow previous studies to measure the industry-time-varying price-cost markups (De Loecker et al. 2020) and the Lerner index (Gutierrez and Philippon 2017) with Compustat data. Since previous studies discuss the methodology in detail, we briefly state how we measure these two variables.

In estimating the price-cost markups, we first use two measures of inputs available in Compustat data, specifically, the cost of goods sold (COGS) and capital, to estimate the output elasticity with respect to the COGS. With the assumption of the firm-level Cobb-Douglas technology, elasticity is the coefficient of the COGS input. We follow Olley and Pakes (1996) for our estimation and allow the elasticity to vary across NACIS 2-digit industries. We leverage the firm-level first-order condition to recover the price-cost markups, which is the estimated elasticity divided by the COGS input share. Then, we aggregate the markups across firms within each 6-digit NAICS industry by using the weight of a sale to recover the industry-level markups. Using a cost-based weight does not make much difference in the results. We measure the Lerner index as the operating income to sales ratio in the 6-digit NAICS industries.

In addition to the two measures of market power, we consider four other control variables to address the industry-time-varying contemporaneous wedge term. Given that standard macroeconomic models predict markup heterogeneity based on price rigidity, we introduce the price rigidity measure from Bils et al. (2013). In light of the literature that links price markup to inventory, we include the inventory-to-sales ratio. We also measure the external financial dependent index of Rajan and Zingales (1998) to control the financial friction that potentially prevents the optimal use of the energy input.

Finally, although the adjustment costs for the energy input are typically assumed away in previous studies, as a robustness check, we integrate the adjustment cost term explicitly. By assuming a quadratic adjustment cost $\Phi(V_{jt}^i) = \frac{\eta^i}{2}\frac{P_{jt}^i}{P_{jt}}V_{jt}^i\left(\frac{V_{jt}^i}{V_{j,t-1}^i} - 1\right)^2$ and solving for the first-order condition (2.5) again, we have

$$\hat{s}^i = \left[\sum_k \delta_{ik}\hat{V}^k\right] - \hat{\phi}^i + \hat{\tau}^i, \tag{B.1}$$

where we have a new term in our equation, $\hat{\phi}_{jt}^i \equiv \eta^i\left(\hat{V}_{jt}^i - \hat{V}_{j,t-1}^i - \frac{E_t[\hat{V}_{j,t+1}^i]-\hat{V}_{jt}^i}{1+\bar{r}}\right)$ with the real interest rate r. Given a small $\bar{r}$ and the (scaled) forecast error $\varepsilon_{j,t+1} \equiv \frac{\eta^i}{1+\bar{r}}(\hat{V}_{j,t+1}^i - E_t[\hat{V}_{j,t+1}^i])$, we have $\hat{\phi}_{jt}^i = \eta^i\left(\hat{V}_{jt}^i - \hat{V}_{j,t-1}^i - \frac{\hat{V}_{j,t+1}^i-\hat{V}_{jt}^i}{1+\bar{r}}\right) + \varepsilon_{j,t+1} \approx -\eta^i\Delta^2\hat{V}_{j,t+1}^i + \varepsilon_{j,t+1}$. We include the double-differences and double-demeaned energy input ($\Delta^2\hat{V}_{j,t+1}^e$) and estimate η^e by instrumenting it with the corresponding lagged input price ($\Delta^2\hat{P}_{j,t-1}^e$) to address the measurement error problem arises

Table B.2: Estimation of Equation (2.6) with Additional Controls

	IV (Lagged Input Prices)					
	(1)	(2)	(3)	(4)	(5)	(6)
production worker	1.261***	1.268***	1.401***	1.517***	1.697***	1.387***
	(0.359)	(0.369)	(0.397)	(0.381)	(0.449)	(0.346)
Markup	-0.158					
	(0.113)					
Lerner index		-69.206**				
		(35.152)				
Frequency of price changes			-0.011**			
			(0.005)			
Inventory				0.497***		
				(0.111)		
RZ measure					-0.006***	
					(0.002)	
Adj. cost						-0.096
						(0.905)
Observations	14591	14591	18198	23217	23220	22275

Note. All columns show the regression result with the instrumental variables. For the instrumental variable analyses, we use both one- and two-year lagged input prices. Five different inputs are used in this regression, specifically, energy, material excluding energy, non-production worker, production worker, and capital. The Markup, Lerner index, Frequency of price changes, RZ measure, and Inventory are controlled directly, and the Adj. cost is instrumented with the corresponding lagged input price. All variables are logged and double-demeaned across industries and time. The standard errors in parentheses are clustered at the NAICS industry level. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

from the forecast error. Because $E_t[\varepsilon_{j,t+1}] = 0$, $\varepsilon_{j,t+1}$ is orthogonal to the instruments and do not violate the exclusion restriction.

Table B.2 shows the empirical results. Regardless of the additional measures of market power, price rigidity, inventory, financial friction, and adjustment costs, our main empirical results of complementarity between production workers and energy input stay positive and statistically significant at the conventional level. Columns (1) and (2) utilize the double-demeaned markup and Lerner index. The coefficients of the measure of market power are intuitive since the market power lowers the input share in a large class of models. Columns (3)-(6) control the measures of price rigidity, inventory-to-sales, financial friction, and adjustment to ease the concerns related to these measures. Consistent with the previous studies, we do not find empirical evidence on the existence of the energy adjustment costs.

We assess the cyclicality of returns to scale conditional on each of these six control variables. We measure the double-demeaned returns to scale by subtracting the double-demeaned wedge term from the double-demeaned energy share, which is a double-demeaned version of Equation (2.9). Figure B.1 shows that the estimated returns to scale still preserves a larger positive correlation with the value added in each industry.

B.3 Returns to Scale: Allowing an Input-Specific Adjustment Cost

In recovering the returns to scale in Section 2.3, we assume that $\tau^i = \tau$. Although this specification is conventional in the previous studies and consistent with the DSGE model presented in Section 3, one concern is that the input-specific adjustment cost is likely to be different across different inputs. This section allows the input-specific adjustment cost and shows that the returns to scale procyclicality is robust to this concern.

Consider the firm's cost minimization problem with the translog production function (2.2):

$$\min C = \sum_i \left[\frac{P^i}{P} V^i + \Phi(V^i) \right]$$

$$s.t.$$

$$\ln(Y) = \ln(z) + \sum_i \beta_i \ln(V^i) + \sum_i \sum_k \frac{\beta_{ik}}{2} \ln(V^i) \ln(V^k) \quad \text{with } \beta_{ik} = \beta_{ki}$$

where $\frac{P^i}{P} \equiv p^i$ is the real input price and $\Phi(V^i)$ is the input-specific adjustment cost for input V^i.

We consider two different functional form assumptions on the adjustment cost $\Phi(V^i)$. First, consider a convex adjustment cost with the following functional form, $\Phi(V^i) \equiv \frac{\eta^i}{2} \frac{P^i}{P} V^i \left(\frac{V^i}{V^i_{ss}} - 1 \right)^2$. In this specification, we make adjustment costs depend on the current variables and steady state values, similar to the specification in Rotemberg and Woodford (1999). Solving the cost-minimization problem with respect to V^i yields

$$s^i \left[1 + \frac{\eta^i}{2} \left(\frac{V^i}{\bar{V}^i} - 1 \right)^2 + \eta^i \left(\frac{V^i}{\bar{V}^i} - 1 \right) \left(\frac{V^i}{\bar{V}^i} \right) \right] = mc \left[\beta_i + \sum_k \beta_{ik} \ln(V^k) \right], \tag{B.2}$$

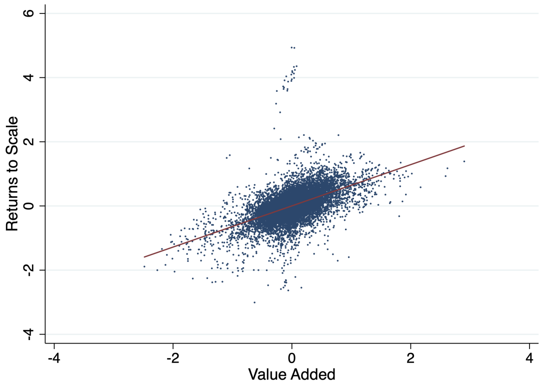
where mc is the Lagrangian multiplier of the production technology, which is the firm's real marginal cost or the inverse price markup. Rearranging and log-linearizing (B.2) leads to

$$\hat{s}^i = \left[\sum_k \delta_{ik} \hat{V}^k \right] + \widehat{mc} - \hat{\phi}^i, \tag{B.3}$$

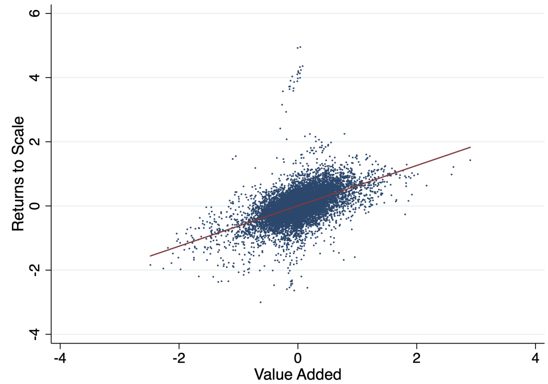
where $\delta_{ik} \equiv \beta_{ik} \left(\frac{\overline{mc}}{\bar{s}^i} \right)$ and $\hat{\phi}^i \equiv \eta^i \hat{V}^i$. Equation (B.3) is a special case of Equation (2.5) with

Figure B.1: Returns to Scale and Value-Added

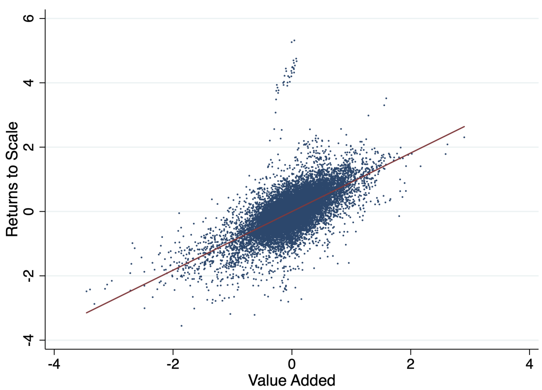
(a) Control Markup



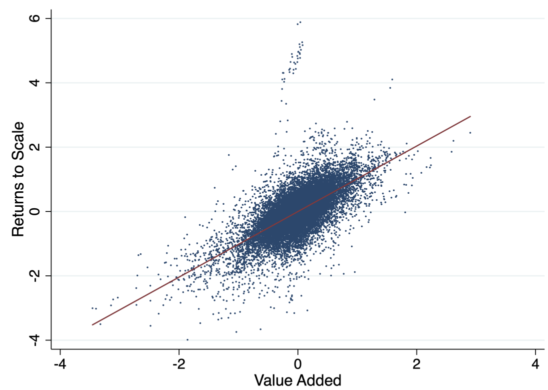
(b) Control Lerner Index



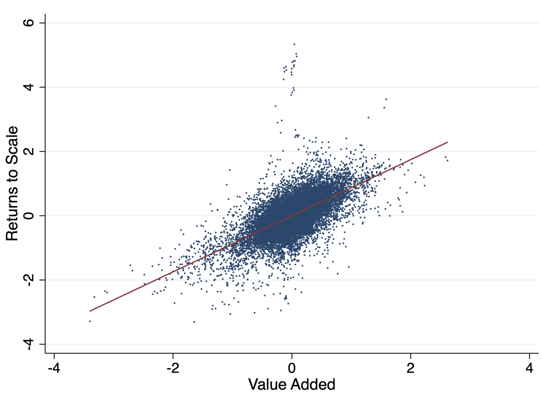
(c) Control Frequency of Price Changes



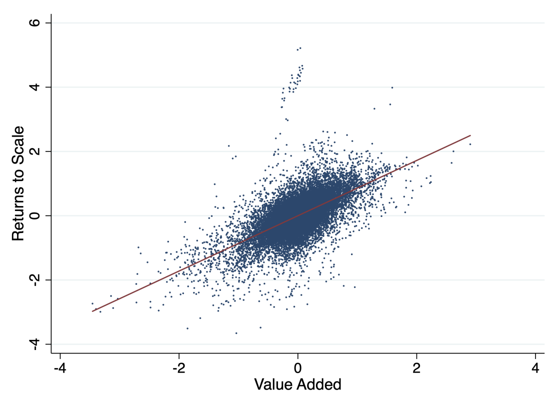
(d) Control RZ Measure



(e) Control the Adjustment Costs



(f) Control Inventory-to-Sales



Note. The y-axis is returns to scale, and the x-axis is value added. All variables are double-demeaned across industries within each year and across years within each industry. The slopes of the linear lines in Figures B.1a, B.1b, B.1c, B.1d, B.1e, and B.1f are .64, .63, .91, 1.02, .87, and .86 respectively.

$\hat{\tau}^i = \widehat{mc} - \hat{\phi}^i$; the input-specific wedge τ^i is the linear combination of the input-specific term arising from the adjustment costs and the inverse price markup in this setup.

By adding Equation (B.2) across all input shares and combining the resulting equation with the returns to scale Equation (2.7), we have

$$rts = \frac{1}{mc} \sum_i s^i \left[1 + \frac{\eta^i}{2} \left(\frac{V^i}{\bar{V}^i} - 1 \right)^2 + \eta^i \left(\frac{V^i}{\bar{V}^i} - 1 \right) \left(\frac{V^i}{\bar{V}^i} \right) \right]. \tag{B.4}$$

The log-linearization of (B.4) leads to

$$\widehat{rts} = \hat{s}^{all} - \widehat{mc} + \hat{\phi}^{all}, \tag{B.5}$$

where $s^{all} = \sum_i s^i$ and $\hat{\phi}^{all} \equiv \sum_i \left(\frac{\bar{s}^i}{\bar{s}^{all}} \right) \hat{\phi}^i = \sum_i \left(\frac{\bar{s}^i}{\bar{s}^{all}} \right) \eta^i \hat{V}^i$.

Equation (B.5) makes it clear that the returns to scale procyclicality results do not likely change with the presence of the input-specific adjustment cost. The key difference between Equation (2.9) and Equation (B.5) is $\hat{\phi}^{all}$, which captures the effects of the input-specific adjustment cost. Equation (2.9) allows the common wedge τ, such as the real marginal cost mc. However, because the input $\hat{V}^i$ is procyclical in the data, regardless of the degree of input-specific adjustment cost η, the adjustment cost term $\hat{\phi}^{all}$ tends to be procyclical. Thus, if adjustment costs exist in any input, then they would strengthen the procyclicality of the returns to scale.

Second, we assume an alternative adjustment cost functional form that depends on lagged inputs, as is the case for the energy input in Appendix B.2: $\Phi(V^i) = \frac{\eta^i}{2} \frac{P_i}{P} V^i \left(\frac{V_t^i}{V_{t-1}^i} - 1 \right)^2$. Under this assumption, one can also derive Equation (B.5) with $\hat{\phi}_t^{all} = \sum_i \frac{\bar{s}^i}{\bar{s}^{all}} \hat{\phi}^i$ with

$$\hat{\phi}_t^i = \eta^i \left(\hat{V}_t^i - \hat{V}_{t-1}^i - \frac{E_t[\hat{V}_{t+1}^i] - \hat{V}_t^i}{1 + \bar{r}} \right),$$

where r is the real interest rate. Since $\left(\hat{V}_t^i - \hat{V}_{t-1}^i - \frac{E_t[\hat{V}_{t+1}^i] - \hat{V}_t^i}{1+\bar{r}} \right) \approx -E_t[\Delta^2 \hat{V}_{t+1}^i]$ with a small $\bar{r}$, the procyclicality of returns to scale might change if $-\Delta^2 \hat{V}_{t+1}^i$ is strongly countercyclical. We assess the cyclcality of $-\Delta^2 \hat{V}_{t+1}^i$ by correlating this measure with the value added for all inputs. We find that $-\Delta^2 \hat{V}_{t+1}^i$ is either procyclical or acyclical and is not likely to weaken the returns to scale procyclicality. The correlation of $-\Delta^2 \hat{V}_{t+1}^i$ with $\hat{Y}_t$ is 0.14, 0.08, 0.0001, 0.12, and 0.07 for production worker, non-production worker, capital, material, and energy, respectively.

B.4 Fixed Costs in Production

This section extends the empirical analyses in Section 2 by allowing the fixed cost of production along with the structural assumptions to be consistent with the DSGE model presented in Section 3.

In the translog production function Equation (2.2), we allow the fixed cost as in Equation (3.1):

$$\ln(Y) = \underbrace{\ln(z) + \sum_i \beta_i \ln(V^i)}_{\text{Cobb-Douglas}} + \underbrace{\sum_i \sum_k \frac{\beta_{ik}}{2} \ln(V^i)\ln(V^k)}_{\text{second-order terms}} - \underbrace{\upsilon}_{\text{fixed costs}} \quad \text{with } \beta_{ik} = \beta_{ki} \tag{B.6}$$

where υ is the fixed cost of production.

As in the DSGE model, we assume that firms face monopolistic competition and earn zero profit at the steady state. Under these assumptions, we derive the first-order condition with respect to input V^i and rearrange the terms to obtain

$$s^i = mc\frac{Y+\upsilon}{Y}\left[\beta_i + \sum_k \beta_{ik}\ln(V^k)\right] \tag{B.7}$$

where mc is the real marginal cost. Equation (B.7) is a special case of Equation (2.4) with $\tau^i = \tau = mc\frac{Y+\upsilon}{Y}$.

As in our main analyses, we choose an energy input as a choice variable ($V^i = E$), allow the variables to change across industry and time, and log-linearize and double-demean Equation (B.7) so that we can implement the estimation. To recover the fixed cost of production in the data, we use the zero-profit condition at the steady state. This condition makes firms use their profit to recover their fixed cost of production at the steady state: $\upsilon = (\Phi - 1)\bar{Y}$, where Φ is the gross price markup at the steady state. Replacing the fixed cost of production, log-linearizing, double-demeaning, and rearranging Equation (B.7) with $V^i = E$ leads to

$$\hat{\hat{s}}^e_{jt} + \frac{\Phi_j - 1}{\Phi_j}\hat{\hat{Y}}_{jt} = \sum_k \delta_{ek}\hat{\hat{V}}^k_{jt} + \widehat{\widehat{mc}}_{jt} \tag{B.8}$$

where Equation (B.8) is a special case of Equation (2.6) with $\hat{\hat{\tau}}^e_{jt} = \widehat{\widehat{mc}}_{jt} - \frac{\Phi_j - 1}{\Phi_j}\hat{\hat{Y}}_{jt}$. The steady state industry-specific value of price markup is recovered from the simple average of markup across years (1958-2009) within the industry, where the industry-time-varying price-cost markup is measured based on De Loecker et al. (2020) methodology from Compustat data as described in Appendix B.2. With the industry-specific price markup measure, Equation (B.8) can be estimated by either treating the double-demeaned marginal cost as a residual or controlling the double-demeaned real marginal cost directly as in Appendix B.2.

Table B.3 reports the results. Regardless of adjusting for the fixed cost of production, we still observe that the coefficient of the production worker, which governs the complementarity between energy and the production worker, is positive and statistically significant. Additionally adjusting for the price markup decreases the coefficient of the production worker but still features large complementarity. The OLS results are largely consistent with the results in Table 1.

Table B.3: Adjusting the Fixed Cost of Production

	IV (Lagged Input Prices)						OLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
energy	-0.849**	-0.449**	-0.861***	-0.176	-0.137	-0.256	0.875***	0.907***
	(0.344)	(0.196)	(0.324)	(0.237)	(0.180)	(0.215)	(0.034)	(0.031)
production worker	2.136***	1.409***	2.106***	1.139***	1.135***	1.248***	-0.394***	-0.420***
	(0.610)	(0.365)	(0.602)	(0.364)	(0.289)	(0.357)	(0.049)	(0.050)
non-production worker	-0.738*	-0.458	-0.689*	-0.691***	-0.653**	-0.591**	-0.115***	-0.136***
	(0.411)	(0.296)	(0.398)	(0.262)	(0.259)	(0.253)	(0.032)	(0.034)
material	0.231	0.639***		0.489***	0.496***		-0.191***	-0.188***
	(0.288)	(0.162)		(0.159)	(0.110)		(0.051)	(0.054)
capital	1.113*		1.402***	0.071		0.733***	0.029	0.060*
	(0.576)		(0.395)	(0.325)		(0.213)	(0.029)	(0.034)
Markup				-0.038	-0.024	-0.069		-0.059
				(0.096)	(0.073)	(0.105)		(0.051)
Observations	20048	20048	20048	14591	14591	14591	20864	14840

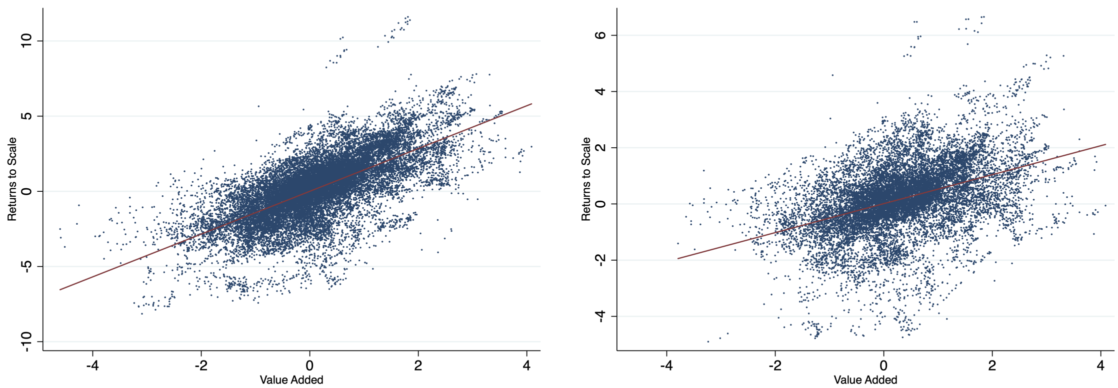
Note: Columns (1)-(3) replicate the regression results in Table 1 columns (1)-(3) by redefining the left hand side variable by subtracting the term relevant to the fixed cost of production ($\frac{\Phi_j - 1}{\Phi_j}\hat{\hat{Y}}_{jt}$) from the energy share. Columns (4)-(6) additionally control the double-demeaned price markup. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

Figure B.2 shows that the procyclicality of returns to scale results remain robust after adjusting for the fixed cost of production. The returns to scale expression (2.9) does not change with the presence of the fixed cost of production except that the wedge term τ becomes the real marginal cost, and there are changes in the estimated production function coefficients used in recovering the returns to scale. Figure B.2a and B.2b are based on Table B.3 columns (1) and (4), respectively. Regardless of the use of alternative specifications, the measure of returns to scale is positively correlated with the value added.

As in our main analyses, separately correlating the components of the returns to scale, which are the price markup and the sum of input shares, still shows the importance of price markup procyclicality in generating the returns to scale procyclicality. Based on Figures B.2a and B.2b, the correlations of value added with returns to scale are 0.64 and 0.5, respectively. The correlations of value added with the price markup are very close to these values—specifically, 0.67 and 0.56. However, the correlation of the sum of input shares and value added is -0.28 and does not correspond to the returns to scale procyclicality.

Figure B.2: Returns to Scale and Value-Added

(a) Adjust Fixed Cost **(b)** Adjust Fixed Cost and Control Markup



Note. The y-axis is returns to scale, and the x-axis is value added. All variables are double-demeaned across industries and year. The slopes of the linear lines in Figures B.2a and B.2b are 1.21 and .59, respectively.

B.5 CES Assumption

This section considers two different methods to test the CES production functional form assumption conditioned on including energy input in the production function. The first method tests the CES assumption with respect to energy input only. The second method tests the CES functional form assumption on any input with a second-order approximation of the production function.

B.5.1 CES Assumption on Energy Input

First, consider the following CES production functional form assumption on the energy input:

$$Q = \big(c(E) + g(V^i; A)\big)^{\frac{1}{\rho}} \tag{B.9}$$

where g(.) is an arbitrary function of other inputs and productivity components. c is an arbitrary constant, and A is productivity. i represents the other four inputs. Rearranging the first-order condition leads to the following equation, similar to Equation (2.4):

$$\log(s^e_{jt}) = \log(c) - \rho\log(Q_{jt}) + \rho\log(V^e_{jt}) + \tau_{jt} \tag{B.10}$$

Note that the CES assumption leads to a testable prediction: the coefficient of output must equal the negative coefficient of energy input. Double-demeaning leads to an equation similar to Equation (2.6):

$$\hat{\hat{s}}^e_{jt} = -\rho\hat{\hat{Q}}_{jt} + \rho\hat{\hat{V}}^e_{jt} + \hat{\hat{\tau}}_{jt} \tag{B.11}$$

Table B.4 estimates Equation (B.11) and rejects the CES assumption on the energy input. Columns (1) and (2) conduct the regression analysis without imposing any constraint. Regardless of the use of instrumental variables or not, the coefficient of output significantly differs from the coefficient of energy input. An instrumental variable (IV) regression result rejects the equality of the absolute value of the coefficients on output and energy with $\chi^2 = 7.82$ (p-value = 0.0052). The OLS results similarly reject the functional form assumption with a near 0 p-value. Columns (3) and (4) re-estimate the coefficients while requiring that the coefficients of output and energy are equal in absolute value. The IV result produces a near-zero coefficient, which supports the Cobb-Douglas functional form assumption with respect to the energy input, which is rejected in our main analysis. The OLS results show that ρ is positive, which does not align with the studies that assume complementarity between energy and other inputs.

Table B.4: Estimation of Equation (2.6) with Additional Controls

	IV	OLS	IV	OLS
	(1)	(2)	(3)	(4)
Output	-0.368***	-0.451***	0.009	-0.497***
	(0.102)	(0.069)	(0.007)	(0.073)
Energy	0.060	0.577***	-0.009	0.497***
	(0.119)	(0.054)	(0.007)	(0.073)
J-test	.1			
(p-value)	.75			
Observations	23694	24167	23221	24167

Note. Columns (1) and (2) present the result based on an unconstrained regression, and columns (3) and (4) present the result based on a constrained regression such that the coefficient of energy input is equal to the negative coefficient of output. Columns (1) and (3) show the regression result with instrumental variables, and columns (2) and (4) show the OLS result. Output, energy, and lagged input prices are logged and double-demeaned across industries and across time. The energy price deflator, total payroll, and price of investment are used to instrument the endogenous variables. Standard errors in parentheses are clustered on the NAICS industry code. J-test refers to Hansen's J-statistics for the overidentifying restriction. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

B.5.2 CES Assumption on Any Input

Second, consider the CES assumption on any input in the production function that explicitly integrates the energy input. We show that assuming the CES assumption requires that $\sum_k \delta_{ik} = 0$ in Equation (2.5). After testing the null hypothesis of $\sum_k \delta_{ik} = 0$ for columns (1), (2), and (3) in Table 1, we strongly reject the null hypothesis with χ^2 of 8.51, 11.55, and 11.89, respectively; the p-values are effectively 0 for all three cases.

Without a loss of generality, we consider below the special cases of a two-factor and three-factor translog production function with energy input for the exposition. It is straightforward to generalize them to five-factor nested CES production functions.

Two-input CES production functions. We begin with a two-factor CES production function. Suppose that

$$y = \exp(\varepsilon^a) \left(\beta_l l^\rho + \beta_e e^\rho\right)^{1/\rho}, \tag{B.12}$$

where $\beta_l + \beta_e = 1$. Two inputs are complementary when $\rho < 0$. A positive ρ implies that two inputs are substitutable. When $\rho = 0$, Equation (B.12) simplifies to a conventional Cobb-Douglas function $\exp(\varepsilon^a) l^{\beta_l} e^{\beta_e}$. In a neighborhood of this Cobb-Douglas function, we approximate Equation (B.12) to the second order as follows.

$$\log y = \underbrace{\varepsilon^a + \beta_l \log l + \beta_e \log e}_{\text{Cobb-Douglas}} + \underbrace{\frac{1}{2}\rho\beta_l\beta_e \left(\log l\right)^2 - \rho\beta_l\beta_e \log l \cdot \log e + \frac{1}{2}\rho\beta_l\beta_e \left(\log e\right)^2}_{\text{second-order terms}} + O(\rho^2). \tag{B.13}$$

Thus, the CES production function in Equation (B.12) includes the second-order terms with tightly parametrized coefficients by ρ, β_l, and β_e. By comparing Equation (B.13) to Equation (2.2), we obtain $\beta_{ll} = \rho\beta_l\beta_e$, $\beta_{el} = -\rho\beta_l\beta_e$, and $\beta_{ee} = \rho\beta_l\beta_e$. As a result, $\beta_{el} + \beta_{ee} = 0$, which further implies that $\delta_{el} + \delta_{ee} = 0$ because $\delta_{ek} = \beta_{ek}\left(\frac{\bar{\tau}^e}{\bar{s}^e}\right)$ for all k (see discussion following Equation (2.5)). We conclude that the above CES production function is compatible with a null hypothesis $\delta_{el} + \delta_{ee} = 0$ that can be tested by using the estimated delta coefficients in Section 2.

Derivation of Equation (B.13). Let $f(\rho; l, e, \varepsilon^a)$ be $\log y = \varepsilon^a + \frac{\log(\beta_l l^\rho + \beta_e e^\rho)}{\rho}$. A Taylor approximation of $\log y$ with respect to ρ around $\rho = 0$ implies $\log y = f(\rho) = f(0) + f'(0)\rho + O(\rho^2)$. Because it is well-known that $f(0) = \varepsilon^a + \beta_l \log l + \beta_e \log e$, we focus on $f'(0)$ here. Note that

$$f'(\rho) = \frac{\frac{\beta_l l^\rho \log l + \beta_e e^\rho \log e}{\beta_l l^\rho + \beta_e e^\rho}\rho - \log\left(\beta_l l^\rho + \beta_e e^\rho\right)}{\rho^2} = \frac{\frac{\beta_l l^\rho \log l + \beta_e e^\rho \log e}{\beta_l l^\rho + \beta_e e^\rho} - \left[f(\rho) - \varepsilon^a\right]}{\rho}.$$

By L'Hospital's rule, we obtain $f'(0) = \left\{ \frac{\beta_l l^\rho \log l + \beta_e e^\rho \log e}{\beta_l l^\rho + \beta_e e^\rho} - [f(\rho) - \varepsilon^a] \right\}' |_{\rho=0}$. As a result, $f'(0) = \frac{1}{2} \left\{ \frac{\beta_l l^\rho \log l + \beta_e e^\rho \log e}{\beta_l l^\rho + \beta_e e^\rho} \right\}' |_{\rho=0}$. Algebraically, we can show that

$$\begin{aligned}
f'(0) &= \lim_{\rho \to 0} \frac{1}{2} \frac{[\beta_l l^\rho (\log l)^2 + \beta_e e^\rho (\log e)^2](\beta_l l^\rho + \beta_e e^\rho) - (\beta_l l^\rho \log l + \beta_e e^\rho \log e)^2}{(\beta_l l^\rho + \beta_e e^\rho)^2} \\
&= \frac{1}{2} \left[\beta_l (\log l)^2 + \beta_e (\log e)^2 - (\beta_l \log l + \beta_e \log e)^2 \right] \\
&= \frac{1}{2} \left[\beta_l (1 - \beta_l)(\log l)^2 - 2\beta_l \beta_e \log l \cdot \log e + \beta_e (1 - \beta_e)(\log e)^2 \right] \\
&= \frac{1}{2} \beta_l \beta_e (\log l - \log e)^2.
\end{aligned}$$

Finally, we obtain the desired result:

$$\log y = f(0) + f'(0)\rho + O(\rho^2) = \varepsilon^a + \beta_l \log l + \beta_e \log e + \frac{1}{2} \rho \beta_l \beta_e (\log l - \log e)^2 + O(\rho^2).$$

Nested CES production functions. Next, we derive similar results for nested CES production functions. We concentrate on nested CES functions with three inputs for exposition. First, we begin with a case where capital is combined with a composite input of labor and energy as follows:

$$y = \exp(\varepsilon^a) \left[\beta_k k^\phi + (\beta_l + \beta_e) g(\rho; l, e)^\phi \right]^{1/\phi}, \tag{B.14}$$

where $g(\rho; l, e) = (\xi_l l^\rho + \xi_e e^\rho)^{1/\rho}$, $\beta_k + \beta_l + \beta_e = 1$, $\xi_l = \frac{\beta_l}{\beta_l + \beta_e}$ and $\xi_e = 1 - \xi_l$. Clearly, Equation (B.14) features constant returns to scale. By repeatedly applying the above result for two-input CES functions, we obtain

$$\begin{aligned}
\log y &= \varepsilon^a + \beta_k \log k + (\beta_l + \beta_e) \log g(\rho) + \frac{1}{2} \phi \beta_k (\beta_l + \beta_e)(\log k - \log g(\rho))^2 + O(\phi^2), \\
\log g(\rho) &= \xi_l \log l + \xi_e \log e + \frac{1}{2} \rho \xi_l \xi_e (\log l - \log e)^2 + O(\rho^2).
\end{aligned}$$

With some algebra, we can show that

$$\begin{aligned}
(\beta_l + \beta_e) \log g(\rho) &= \beta_l \log l + \beta_e \log e + \frac{1}{2} \frac{\rho \beta_l \beta_e}{\beta_l + \beta_e} (\log l - \log e)^2 + O(\rho^2), \\
\phi(\log k - \log g(\rho))^2 &= \phi \left[\log k - \xi_l \log l - \xi_e \log e + O(\rho) \right]^2 = \phi(\log k - \xi_l \log l - \xi_e \log e)^2 + O(\phi\rho) + O(\phi\rho^2).
\end{aligned}$$

Thus, we have the following translog approximation of Equation (B.14):

$$\begin{aligned}
\log y &= \varepsilon^a + \beta_k \log k + \beta_l \log l + \beta_e \log e \\
&\quad + \frac{1}{2} \frac{\rho \beta_l \beta_e}{\beta_l + \beta_e} (\log l - \log e)^2 + \frac{1}{2} \phi \beta_k (\beta_l + \beta_e)(\log k - \xi_l \log l - \xi_e \log e)^2 + O(||\phi, \rho||^2),
\end{aligned} \tag{B.15}$$

which implies that $\beta_{ek} = -\phi\beta_k(\beta_l+\beta_e)\xi_e$, $\beta_{el} = -\rho\frac{\beta_l\beta_e}{\beta_l+\beta_e}+\phi\beta_k(\beta_l+\beta_e)\xi_l\xi_e$ and $\beta_{ee} = \rho\frac{\beta_l\beta_e}{\beta_l+\beta_e}+\phi\beta_k(\beta_l+\beta_e)\xi_e^2$. We conclude that $\beta_{ek}+\beta_{el}+\beta_{ee}=0$, and therefore, $\delta_{ek}+\delta_{el}+\delta_{ee} = (\beta_{ek}+\beta_{el}+\beta_{ee})\frac{\bar{\tau}^e}{\bar{s}^e} = 0$.

We can obtain a similar result for the case where labor is combined with a composite input of capital and energy, $y = \exp(\varepsilon^a)\left[\beta_l l^\phi + (\beta_k+\beta_e)g(\rho;k,e)^\phi\right]^{1/\phi}$, by switching the roles between k and l above. Finally, we investigate the case where energy is combined with a composite input of capital and labor: $y = \exp(\varepsilon^a)\left[\beta_e e^\phi + (\beta_k+\beta_l)h(\rho;k,l)^\phi\right]^{1/\phi}$, where $h(\rho;k,l) = (\zeta_k k^\rho + \zeta_l l^\rho)^{1/\rho}$, $\zeta_k = \frac{\beta_k}{\beta_k+\beta_l}$ and $\zeta_l = 1-\zeta_k$. In this case, we have

$$\log y = \varepsilon^a + \beta_e \log e + (\beta_k+\beta_l)\log h(\rho) + \frac{1}{2}\phi\beta_e(\beta_k+\beta_l)(\log e - \log h(\rho))^2 + O(\phi^2),$$

$$\log h(\rho) = \zeta_k \log k + \zeta_l \log l + \frac{1}{2}\rho\zeta_k\zeta_l(\log k - \log l)^2 + O(\rho^2),$$

$$(\beta_k+\beta_l)\log h(\rho) = \beta_k \log k + \beta_l \log l + \frac{1}{2}\frac{\rho\beta_k\beta_l}{\beta_k+\beta_l}(\log k - \log l)^2 + O(\rho^2),$$

$$\phi(\log e - \log h(\rho))^2 = \phi\left[\log e - \zeta_k \log k - \zeta_l \log l + O(\rho)\right]^2 = \phi(\log e - \zeta_k\log k - \zeta_l \log l)^2 + O(\phi\rho) + O(\phi\rho^2).$$

Thus, we have the following translog approximation of this nested CES production function:

$$\log y = \varepsilon^a + \beta_k \log k + \beta_l \log l + \beta_e \log e + \frac{1}{2}\frac{\rho\beta_k\beta_l}{\beta_k+\beta_l}(\log k - \log l)^2 + \frac{1}{2}\phi\beta_e(\beta_k+\beta_l)(\log e - \zeta_k\log k - \zeta_l\log l)^2 + O(||\phi,\rho||^2).$$

Here, $\beta_{ek} = -\phi\beta_e(\beta_k+\beta_l)\zeta_k$, $\beta_{el} = -\phi\beta_e(\beta_k+\beta_l)\zeta_l$ and $\beta_{ee} = \phi\beta_e(\beta_k+\beta_l)$. Again, we have $\beta_{ek}+\beta_{el}+\beta_{ee}=0$.

Accordingly, for any three-input nested CES production function, we have $\beta_{ek}+\beta_{el}+\beta_{ee} = \delta_{ek}+\delta_{el}+\delta_{ee} = 0$.

B.6 Regression Weights

This section revisits the main empirical analyses reported in Table 1 by weighting different industries so that the estimated parameter better matches the aggregate parameter. We consider both the industry-time-varying weight and industry-specific weights. As shown in Table B.5, the complementarity between labor and energy is preserved with the weighting; the corresponding parameter ranges from 1.26 to 1.51 depending on the weights.

B.7 Smaller Number of Inputs

The empirical analyses in Section 2 utilize five inputs (energy, capital, production worker, non-production worker, and material) available in the NBER-CES data, whereas the DSGE model presented in Section 3 allows for only three inputs (energy, capital, and labor), which generates a

Table B.5: Weighted Regression

	IV (Lagged Input Prices)						OLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
energy	-0.195	-0.212	-0.288	-0.635***	-0.454***	-0.630***	0.792***	0.885***
	(0.200)	(0.250)	(0.231)	(0.214)	(0.172)	(0.209)	(0.067)	(0.043)
production worker	1.259***	1.302***	1.533***	1.515***	1.157***	1.565***	-0.272***	-0.334***
	(0.419)	(0.383)	(0.475)	(0.382)	(0.225)	(0.358)	(0.098)	(0.080)
non-production worker	-0.496	-0.496*	-0.670	-0.488	-0.137	-0.525	-0.081	-0.077
	(0.382)	(0.267)	(0.444)	(0.324)	(0.190)	(0.356)	(0.071)	(0.064)
material	-0.126	-0.114		-0.227	-0.133		-0.438***	-0.438***
	(0.102)	(0.099)		(0.150)	(0.091)		(0.054)	(0.053)
capital	0.011		0.008	0.534		0.301	-0.033	-0.044
	(0.242)		(0.271)	(0.372)		(0.300)	(0.072)	(0.055)
J-test	11.77	11.85	14.78	12.02	14.76	14.8		
(p-value)	.04	.07	.02	.03	.02	.02		
Observations	23220	23220	23220	23220	23220	23220	23220	23220

Note: To inform on the aggregate parameters, we weight the observations by either industry-time-specific output ("industry-time-weight") or industry-specific output ("industry weight"). Other regression specifications are identical to those in Table 1. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

discrepancy between the model and the data. This section revisits the empirical analyses and allows a smaller number of inputs in the estimation, consistent with the DSGE model.

Table B.6 shows that the strong complementary between labor and energy remains regardless of combining production and non-production workers or using a value-added production function that excludes material input. Column (1) combines both production worker and non-production into all workers, column (2) excludes material input, thereby considering the value-added production function, and column (3) considers both the combined worker input and the value-added production function. Regardless of using different specifications, we still estimate the complementarity between labor and energy that is within a range of the Bayesian estimate in Section 3. The OLS results also show a similar empirical pattern as in Table 1; the estimated coefficient of energy input is positive, and the other coefficients are negative.

B.8 Using the KLEMS Data

One potential concern in using the NBER-CES database is that it only covers the manufacturing sectors, which does not represent the entire US economy. Thus, in introducing the complementarity parameter into the DSGE model to understand the macroeconomic implications, we use the estimated parameter from the NBER-CES database as a prior and re-estimate the parameter with the aggregate data.

As an additional robustness exercise, this section analyzes the KLEMS database. Unlike the

Table B.6: Estimation with a Smaller Number of Inputs

	IV (Lagged Input Prices)			OLS		
	(1)	(2)	(3)	(4)	(5)	(6)
energy	-0.347	-0.854**	-0.929**	0.852***	0.789***	0.794***
	(0.231)	(0.392)	(0.437)	(0.029)	(0.034)	(0.034)
all workers	1.347***		1.858***	-0.458***		-0.692***
	(0.395)		(0.639)	(0.033)		(0.032)
production worker		1.848***			-0.458***	
		(0.626)			(0.035)	
non-production worker		-0.450			-0.228***	
		(0.355)			(0.028)	
material	0.257			-0.400***		
	(0.227)			(0.027)		
capital	-0.190	0.854*	0.376	-0.024	-0.203***	-0.187***
	(0.370)	(0.497)	(0.348)	(0.021)	(0.024)	(0.024)
value-added share		✓	✓		✓	✓
Observations	23221	23220	23221	24167	24166	24167

Note: The "all workers" add the production worker and non-production worker. When material input is excluded in the regression, we use the value added for the energy share instead of the value of shipment. The energy share is defined as the energy expenditure over the value added, which subtracts material expenditure from the value of shipment. In using the instruments, only the input prices that correspond to included inputs are used. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

NBER-CES database, the KLEMS database covers the entire sector of the US economy. It includes key inputs that we use in this paper, such as energy, labor, capital, and material inputs, as well as its price indexes. It is balanced industry panel data that cover 60 industries over 26 years from 1987 to 2012. This database is available on the Bureau of Economic Analysis (BEA) website and is used by influential papers, such as Bils et al. (2018). Although the data cover aggregate sectors over a short period relative to the NBER-CES database, to the best of our knowledge, it is the only industry-level panel data that have all the relevant information for our estimation technique and cover the entire US economy.

Table B.7: Using the KLEMS Data

	IV (Lagged Input Prices)						OLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
energy	1.369***	1.245***	1.214***	1.439***	1.182***	1.314***	0.906***	0.898***
	(0.307)	(0.318)	(0.232)	(0.301)	(0.278)	(0.218)	(0.111)	(0.090)
all workers	1.578**	1.674***	1.489***	1.574**	1.253**	1.164**	-0.090	-0.210*
	(0.658)	(0.618)	(0.506)	(0.667)	(0.525)	(0.521)	(0.130)	(0.110)
material	-0.064	0.048		-0.071	0.112		-0.181**	-0.181***
	(0.292)	(0.278)		(0.269)	(0.244)		(0.069)	(0.060)
capital	0.165		0.215	0.254		0.338	0.043	0.025
	(0.302)		(0.276)	(0.306)		(0.277)	(0.074)	(0.066)
Industry-time weight	✓	✓	✓				✓	
Industry weight				✓	✓	✓		✓
Observations	1020	1020	1020	1020	1020	1020	1140	1140

Note: The "all workers" is the labor input available in the KLEMS data. To be consistent with the exercise that uses the NBER-CES data, we define the energy share after subtracting the service input from the output. To inform on the aggregate parameters, we weight the observations by either industry-time-specific output ("industry-time-weight") or industry-specific output ("industry weight"). As in Table 1, both the one- and two-year lagged input prices are used as instrumental variables. *, **, and *** indicate significance at the 0.1, 0.05, and 0.01 levels, respectively.

We re-estimate the parameters based on Equation (2.6) with the KLEMS data and replicate Table 1. Table B.7 presents the results. Given that the motivation of this exercise is to inform on the aggregate representative parameters, we weight the sample either by industry-time-specific output (columns 1-3 and 7) or industry-specific output (columns 4-6 and 8). Regardless of the use of different specifications, we still find a strong complementarity between labor and energy similar to our main empirical analysis in Table 1. In addition, the complementarity parameter between labor and energy is in the range with the Bayesian estimation results in Section 3.2. One difference from Table 1 is that the energy input no longer features the strong diminishing returns, potentially because industries outside of the manufacturing sectors do not utilize the energy input the same as in the manufacturing sectors. The OLS results show a similar empirical pattern.

Appendix C Translog Production Function in the DSGE Model

C.1 Properties of the Translog Production Function

In this section, we discuss in more detail the properties of the translog production function that we propose. For the clarity of exposition, we abstract away fixed cost (i.e., $v = 0$) from Equation (3.1):

$$y_t(i) = \exp(\varepsilon_t^a)[k_t^s(i)]^{\beta_k}[l_t(i)]^{\beta_l}[e_t(i)]^{\beta_e} \times \left[\left(\frac{l_t(i)}{\bar{l}}\right)^{\beta_{el}\log(e_t/\bar{e})}\left(\frac{e_t(i)}{\bar{e}}\right)^{\beta_{el}\log(l_t/l)}\right]. \tag{C.1}$$

Equation (C.1) can be broken down into the short-run and long-run components, in which the short-run component is expressed as the deviation from the long-run component (i.e., the steady state):

$$y_t(i) = f^{SR}\left(\frac{k_t^s(i)}{\bar{k}}, \frac{l_t(i)}{\bar{l}}, \frac{e_t(i)}{\bar{e}}; \Omega_t\right) \cdot f^{LR}\left(\bar{k}, \bar{l}, \bar{e}; \Omega\right), \tag{C.2}$$

where

$$\begin{aligned} f^{SR}\left(\frac{k_t(i)}{\bar{k}}, \frac{l_t(i)}{\bar{l}}, \frac{e_t(i)}{\bar{e}}; \Omega_t\right) &\equiv \exp(\varepsilon_t^a)\left(\frac{k_t^s(i)}{\bar{k^s}}\right)^{\beta_k}\left(\frac{l_t(i)}{\bar{l}}\right)^{\beta_l+\beta_{el}\log(e_t/\bar{e})}\left(\frac{e_t(i)}{\bar{e}}\right)^{\beta_e+\beta_{el}\log(l_t/\bar{l})} \\ &\left(= \frac{y_t(i)}{\bar{y}}\right), \\ f^{LR}(\bar{k^s}, \bar{l}, \bar{e}; \Omega) \equiv \left(\bar{k^s}\right)^{\beta_k}\left(\bar{l}\right)^{\beta_l}\left(\bar{e}\right)^{\beta_e} (= \bar{y}) \end{aligned} \tag{C.3}$$

with

$$\begin{aligned} &\Omega_t \equiv [\beta_k, \beta_l + \beta_{el}\log(e_t/\bar{e}), \beta_e + \beta_{el}\log(l_t/\bar{l})]', \\ &\Omega \equiv [\beta_k, \beta_l, \beta_e]'. \end{aligned} \tag{C.4}$$

A similar decomposition can be found in Cantore et al. (2015) and Koh and Santaeulàlia-Llopis (2017) in the context of the CES production function in an attempt to resolve dimensionality issues Cantore and Levine (2012).[23] Here, l_t and e_t indicate the cross-sectional average of $l_t(i)$ and $e_t(i)$, respectively, which individual firms take as a given.

In the above expression, Ω_t is a vector of *endogenous* parameters that govern potentially time-

[23] As discussed in De Jong (1967), Cantore and Levine (2012) and Cantore et al. (2014), such a distinction between short- and long-run production functions (or normalization) was implemented to address what is called the dimensionality issue: under CES technology, factor share parameters no longer directly measure the "share" but depend on the underlying "dimensions". This makes such parameters depend on the choice of units, which creates measurement problems in calibration and estimation. Expressing the production function in deviation from the steady-state eliminates the dimensional parameters and resolves the issue. As discussed below, our normalization also makes the translog parameters under our production function no longer depend on the choice of units.

varying short-run returns to scale of the economy. We write "endogenous" parameters because the time-varying components are endogenously determined in equilibrium, although individual firms take these values as a given.[24] Ω is a vector of strictly exogenous parameters that govern the time-invariant and constant long-run returns to scale of the economy. If we evaluate Ω_t at the steady-state or on the long-run horizon, we have $\overline{\Omega} = \Omega$.

The proposed translog production function has the following properties.

1. The bracketed term on the right-hand side of (C.1), $\left[\left(\frac{l_t(i)}{\bar{l}}\right)^{\beta_{el}\log(e_t/\bar{e})}\left(\frac{e_t(i)}{\bar{e}}\right)^{\beta_{el}\log(l_t/l)}\right]$ consists of the variables normalized by their steady state values. This makes the production function collapse into the conventional Cobb-Douglas at the steady state.

   There are three advantages of the normalization. First, it resolves the dimensionality issue discussed in De Jong (1967), Cantore and Levine (2012) and Cantore et al. (2014); that is, the normalization makes β_{el} independent of the choice of units. To illustrate this point, note that under the normalization, the labor and energy input shares are given by $\frac{w_t l_t}{y_t} = \lambda\left[\beta_l + \beta_{el}\log\left(e_t/\bar{e}\right)\right]$ and $\frac{p_t^e e_t}{y_t} = \lambda\left[\beta_e + \beta_{el}\log\left(l_t/\bar{l}\right)\right]$, respectively, where λ is the wedge between the marginal product of inputs and real input prices. From this expression, we can see that the parameter β_{el} is dimensionless because the inputs that appear on the right-hand sides are normalized by their steady state counterparts (i.e., $\frac{e_t}{\bar{e}}$ and $\frac{l_t}{\bar{l}}$), which do not depend on the choice of units. In contrast, if we do not normalize the production function, then the labor and energy input shares become $\frac{w_t l_t}{y_t} = \lambda\left[\beta_l + \beta_{el}\log e_t\right]$ and $\frac{p_t^e e_t}{y_t} = \lambda\left[\beta_e + \beta_{el}\log l_t\right]$, respectively. In this case, since the input shares are dimensionless, β_{el} depends on the unit of inputs (e_t and l_t).

   Second, it makes the model compatible with the balanced-growth path.[25] In this way, the input complementarity that we introduce (and the resulting procyclical returns to scale) is a short-run characteristic, which does not affect the long-run growth of the economy.

   Finally, the normalization facilitates a comparison with the model without complementarity-induced procyclical returns to scale because the steady-state is identical across the two models.

2. At the aggregate level, the short-run component f^{SR} has the translog expression:

$$\log\left(\frac{y_t}{\bar{y}}\right) = \varepsilon_t^a + \beta_k\log\left(\frac{k_t^s}{\bar{k}^s}\right) + \beta_l\log\left(\frac{l_t}{\bar{l}}\right) + \beta_e\log\left(\frac{e_t}{\bar{e}}\right) + 2\beta_{el}\log\left(\frac{l_t}{\bar{l}}\right)\log\left(\frac{e_t}{\bar{e}}\right)$$

3. The complementarity between energy and labor is reflected by a single parameter β_{el}. If $\beta_{el} > 0$, our model features complementarity-induced procyclical returns to scale in the short-run,

[24] Because the time-varying components in Ω_t are endogenously determined in equilibrium, the terminology "parameter" can be somewhat misleading. Still, we call it an endogenous parameter because it characterizes the returns to scale of the economy that individual firms take as exogenous. The long-run value of Ω_t, Ω, is a vector of strictly exogenous parameters in the sense that it consists only of deep structural parameters.

[25] Our medium-scale DSGE model features the balanced-growth path.

provided that the dynamics of $\log(e_t/\bar{e})$ and $\log(l_t/\bar{l})$ are procyclical.

4. Despite procyclical returns to scale, the production function becomes *scale-free* up to the first order because log-linearizing the production function yields a form exactly identical to the log-linearized Cobb-Douglas production function. To see this, consider the first-order approximation of the production function at the aggregate level:

$$\hat{y}_t = \varepsilon_t^a + \beta_k \hat{k}_t^s + \beta_l \hat{l}_t + \beta_e \hat{e}_t,$$

where $\hat{x}_t \equiv \log \frac{x_t}{\bar{x}}$ for an arbitrary variable x.

Therefore, the procyclicality of returns to scale does not generate any additional fluctuation of output by itself and behaves exactly the same as the conventional Cobb-Douglas up to the first order. All interesting dynamics arise through the first-order condition of the firm. This *scale-free* characteristic up to the first order is one feature that distinguishes our model from the conventional increasing returns to scale model. In the increasing returns to scale model, both the production function and the first-order conditions are affected by the increasing returns up to the first order, while only the first-order conditions are affected by the procyclical returns to scale in our model.

5. The short- and long-run returns to scale at the individual firm level are given by

$$\begin{aligned} &\text{Short-Run Returns to Scale } (rts_t) = \beta_k + \beta_l + \beta_e + \beta_{el}\left[\log(e_t/\bar{e}) + \log(l_t/\bar{l})\right] \\ &\text{Long-Run Returns to Scale } (\overline{rts}) = \beta_k + \beta_l + \beta_e \end{aligned} \tag{C.5}$$

Since individual firms take the cross-sectional average variables e_t and l_t as given and do not internalize their changes, each individual firm takes the returns to scale as a given.

This assumption guarantees that firms' optimizing behavior is well-characterized by the first-order conditions. If an individual firm can internalize the change in returns to scale of the economy, then by choosing a larger amount of labor and energy inputs, each firm can make their returns to scale arbitrarily large. This would induce firms to choose an infinite amount of labor and energy inputs. By assuming that individual firms do not internalize the change in returns to scale, this issue no longer arises.[26]

[26] This assumption makes our model similar to the internal increasing returns to scale (IRS) model. In contrast to an external IRS model in which individual firms take the production externalities as a given and therefore effectively face constant returns to scale, firms in the internal IRS model take the returns to scale parameter (which is larger than one) as a given and therefore face increasing returns to scale. In our model, individual firms do not internalize the change in returns to scale, and firms take as a given the returns to scale evaluated at the aggregate inputs.

C.2 Aggregation of the Firm-level Production Function

Suppose that X_t is an aggregate quantity of firm-level variable $X_t(i)$ such that $X_t = \int X_t(i)di$. We denote the logarithm of X_t and $X_t(i)$ by x_t and $x_t(i)$, respectively.

Lemma. $x_t = \int x_t(i)\ di$ up to the first order.

Proof. We closely follow the discussion in Galí (2015, Appendix 3.4).

Let $E_i\{x_t(i)\}$ be $\int x_t(i)\ di$. Similarly, we define $var_i\{x_t(i)\}$ as $\int(x_t(i) - E_i\{x_t(i)\})^2\ di$.

From the definition of X_t,

$$\begin{aligned} 1 &= \int \frac{X_t(i)}{X_t}\ di = \int \exp(x_t(i) - x_t)\ di \\ &\approx 1 + \int (x_t(i) - x_t)\ di + \frac{1}{2}\int (x_t(i) - x_t)^2\ di, \end{aligned}$$

where the last line is based on the second-order Taylor expansion. Thus, we have

$$x_t \approx E_i\{x_t(i)\} + \frac{1}{2}\int (x_t(i) - x_t)^2\ di$$

up to the second order.

It remains to show that $A_t \equiv \int(x_t(i) - x_t)^2\ di$ is of the second-order. The above equation implies that $E_i\{x_t(i)\} - x_t \approx -\frac{1}{2}A_t$ up to the second order. Thus,

$$\begin{aligned} A_t &= \int (x_t(i) - E_i\{x_t(i)\} + E_i\{x_t(i)\} - x_t)^2\ di \\ &= \int (x_t(i) - E_i\{x_t(i)\})^2\ di + 2\int (x_t(i) - E_i\{x_t(i)\})(E_i\{x_t(i)\} - x_t)\ di + (E_i\{x_t(i)\} - x_t)^2 \\ &\approx var_i\{x_t(i)\} - A_t \int (x_t(i) - E_t\{x_t(i)\})\ di + \left(\frac{1}{2}A_t\right)^2 \\ &= var_i\{x_t(i)\} + \left(\frac{1}{2}A_t\right)^2 \end{aligned}$$

up to the second order. This result implies that

$$A_t \approx 2 \pm 2\sqrt{1 - var_i\{x_t(i)\}}$$

up to the second order. Because $\sqrt{1 - var_i\{x_t(i)\}} \approx 1 - \frac{var_i\{x_t(i)\}}{2} - \frac{[var_i\{x_t(i)\}]^2}{8}$ up to the second order,

$$\begin{aligned} A_t &\approx 2 \pm 2\sqrt{1 - var_i\{x_t(i)\}} \\ &\approx 2 \pm 2\left(1 - \frac{var_i\{x_t(i)\}}{2} - \frac{[var_i\{x_t(i)\}]^2}{8}\right). \end{aligned}$$

We know that $A_t > 0$. Thus, we obtain

$$\begin{aligned} A_t &\approx 2 - 2\left(1 - \frac{var_i\{x_t(i)\}}{2} - \frac{[var_i\{x_t(i)\}]^2}{8}\right) \\ &\approx var_i\{x_t(i)\} \end{aligned}$$

up to the second order. Because $var_i\{x_t(i)\}$ is of the second order, the proof is complete. □

Consider the firm-level production function (3.1). This implies that

$$\log(y_t(i) + \upsilon) = \varepsilon_t^a + \beta_k \log(k_t^s(i)) + \beta_l \log(l_t(i)) + \beta_e \log(e_t(i)) + \beta_{el}\hat{e}_t\hat{l}_t(i) + \beta_{el}\hat{e}_t(i)\hat{l}_t.$$

For the Kimball aggregator, it is well-known that $y_t \approx \int y_t(i)\, di$ up to the first order (Smets and Wouters, 2007). Thus, by repeatedly applying the above lemma, we have

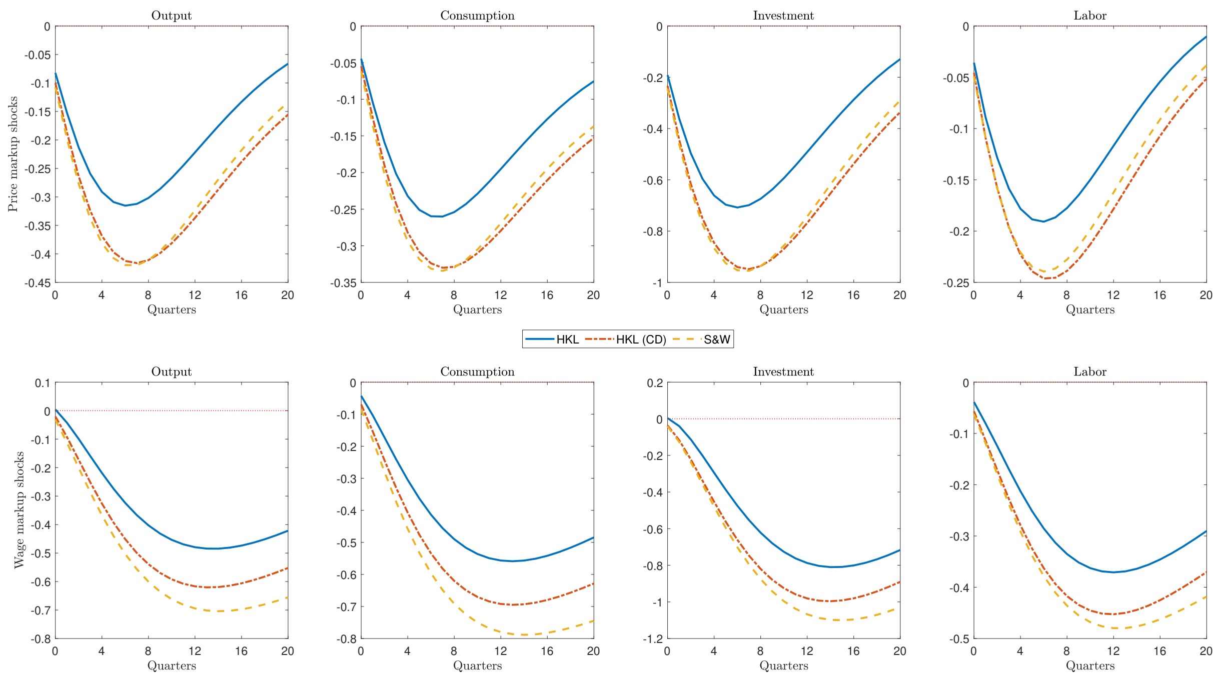
$$\begin{aligned} \log(y_t + \upsilon) &\approx \log\left(\int y_t(i)\, di + \upsilon\right) \approx \int \log(y_t(i) + \upsilon)\, di \\ &\approx \varepsilon_t^a + \beta_k \log k_t^s + \beta_l \log l_t + \beta_e \log e_t + \beta_{el}\hat{e}_t \left[\int \log(l_t(i))\, di - \log \bar{l}\right] \\ &+ \beta_{el}\left[\int \log e_t(i)\, di - \log \bar{e}\right]\hat{l}_t \\ &\approx \varepsilon_t^a + \beta_k \log k_t^s + \beta_l \log l_t + \beta_e \log e_t + \beta_{el}\hat{e}_t \log\left(\frac{l_t}{\bar{l}}\right) + \beta_{el}\log\left(\frac{e_t}{\bar{e}}\right)\hat{l}_t \\ &= \varepsilon_t^a + \beta_k \log k_t^s + \beta_l \log l_t + \beta_e \log e_t + 2\beta_{el}\hat{l}_t\hat{e}_t, \end{aligned}$$

where all the approximation is up to the first order. By taking the exponential, we recover the aggregate translog production function that mirrors the firm-level production function up to the first order.

$$y_t \approx \exp(\varepsilon_t^a)[k_t^s]^{\beta_k}[l_t]^{\beta_l}[e_t]^{\beta_e}\left(\frac{l_t}{\bar{l}}\right)^{\beta_{el}\log(e_t/\bar{e})}\left(\frac{e_t}{\bar{e}}\right)^{\beta_{el}\log(l_t/\bar{l})} - \upsilon.$$

Appendix D Additional Impulse Responses

Figure D.1: Impulse responses to price and wage markup shocks



Notes. The top panels show the responses of output, consumption, investment, and labor to a one-standard-deviation contractionary price markup shock. The solid, dashdotted, and dashed lines represent the results based on our benchmark model (HKL), the model without input complementarity (HKL-CD), and the Smets and Wouters (2007) model (S&W), respectively. Similar responses to a one-standard-deviation contractionary wage markup shock are illustrated in the bottom panels.

Figure D.1 shows the impulse responses of major macroeconomic variables in response to price and wage markup shocks at the posterior mode of each of the three models. The top panels illustrate the responses of output, consumption, investment, and labor to a one-standard-deviation contractionary price markup shock. The solid, dashdotted, and dashed lines represent the results based on our benchmark model (HKL), the model without input complementarity (HKL-CD), and the Smets and Wouters (2007) model (S&W), respectively. Similar responses to a one-standard-deviation contractionary wage markup shock are illustrated in the bottom panels.

Clearly, HKL features the smallest responses of major aggregate variables to price and wage markup shocks. For instance, the peak effects on output of a one-standard-deviation price markup shock is 0.32%, 0.42%, and 0.42% at quarterly frequencies based on HKL, HKL-CD, and S&W, respectively. For wage markup shocks, the peak effects on output are 0.48%, 0.62%, and 0.70% at quarterly frequencies. The results for consumption, investment, and labor are similar.