

Agnostic Structural Disturbances (ASDs): Detecting and Reducing Misspecification in Empirical Macroeconomic Models

Wouter J. Den Haan, Thomas Drechsel

Abstract

Constructing empirical specifications for structural economic models is difficult, if not impossible. As shown in this paper, even minor misspecifications may lead to large distortions for parameter estimates and implied model properties. We propose a novel concept, namely an agnostic structural disturbance (ASD), that can be used to both detect and correct for misspecification of structural disturbances and is easy to implement. While agnostic in nature, the estimated coefficients and associated impulse response functions of these ASDs allow us to give them an economic interpretation. We adopt the methodology to the Smets-Wouters model and formulate an improved risk-premium and an improved investment-specific productivity disturbance.

Keywords: DSGE, full-information model estimation, structural disturbances

1. Introduction

Exogenous random shocks are the lifeblood of modern macroeconomic business cycle models. They enter the model as innovations to structural disturbances that affect key aspects of the model. Recent generations of business cycle models include a multitude of structural disturbances. Structural disturbances impose restrictions on model equations and,

*Den Haan: Centre for Macroeconomics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK and CEPR, London, UK. E-mail: wjdenhaan@gmail.com. Drechsel: Centre for Macroeconomics, London School of Economics and Political Science. We would like to thank Fabio Canova, Jesus Fernandez-Villaverde, Pablo Guerron-Quintana, Marek Jarociński, Christian Matthes, Kris Nimark, Giorgio Primiceri, Ricardo Reis, Barbara Rossi, Luca Sala, Frank Schorfheide, and Antonella Trigari for useful comments. This research is supported by the Economic and Social Research Council (ESRC).

6 thus, on the model’s solutions. Therefore, each structural disturbance has to enter *each*
7 model equation correctly. This is a concern, since we often do not have independent evidence
8 on how structural disturbances should affect the system. For example, should a risk-premium
9 disturbance affect all Euler equations or only some? Is it correct to assume that structural
10 disturbances are uncorrelated as is commonly done? Chari et al. (2007) propose “wedges”
11 as alternatives to standard structural disturbances, but wedges also impose restrictions.

12 The contributions of this paper are threefold. First, we propose the agnostic structural
13 disturbance (ASD) as an alternative type of *structural* disturbance. The procedure simply in-
14 volves adding structural disturbances with associated reduced-form coefficients to *each* model
15 equation or alternatively to each policy rule. In contrast to regular structural disturbances
16 and wedges, ASDs impose *no* additional restrictions on policy rules. Nevertheless, they are
17 different from measurement error, because they are structural and propagate through the
18 system like regular structural disturbances. The procedure of Cúrdia and Reis (2012) shares
19 with ours the ability to deal with correlated structural disturbances, but their disturbances
20 still impose all the restrictions on model equations of regular structural disturbances.

21 Our ASD procedure can be used to test whether regular structural disturbances are cor-
22 rectly specified and to enrich an empirical specification by adding ASDs as additional struc-
23 tural disturbances. Using Monte Carlo experiments, we document that the ASD procedure
24 is capable of detecting and correcting for misspecification in samples of typical size.

25 The second contribution of our paper is to test whether the structural disturbances of the
26 model in Smets and Wouters (2007) (SW) are correctly specified using the same US postwar
27 data set. We find that the risk-premium and the investment-specific productivity disturbance
28 are not. We use our procedure to improve on the SW empirical specification. Our preferred
29 specification (based on marginal likelihood considerations) has three ASDs and excludes the
30 SW risk-premium and the SW investment-specific disturbance.

31 Although the ASD procedure itself does not rely on any economic reasoning, the estima-
32 tion results – both the associated coefficients and their impulse response functions (IRFs)
33 – may reveal a lot about the type of structural disturbance the data has identified. For
34 example, we interpret one of the ASDs in our adjusted empirical specification of the SW

35 model as an “investment-modernization” disturbance, because it stimulates investment, but
 36 at the same time leads to faster depreciation of the existing capital stock. The second ASD
 37 of our empirical model has features in common with both a risk-premium and a preference
 38 disturbance but is also different from both. Finally, the third ASD captures increases in the
 39 wage mark-up disturbance that are associated with an increase in the utilized capital stock.

40 The third contribution of our paper consists of showing that *minor* misspecifications of
 41 the empirical model regarding structural disturbances can easily lead to *large* distortions for
 42 parameter estimates and model properties, such as business cycle statistics and IRFs. We
 43 document that ASDs can alleviate these problems.

44 The next section explains the ASD procedure. Section 3 documents the ability of ASDs
 45 to detect and correct for misspecification using Monte Carlo experiments for a typical appli-
 46 cation. Section 4 discusses the results when our procedure is applied to the SW model on
 47 US data. Section 5 concludes.

48 2. Agnostic Structural Disturbances

49 We use a simple business cycle model to explain what ASDs are and how they can be used
 50 for building theoretical models that one wants to bring to the data. Appendix A, provides
 51 a general formulation.

52 2.1. Model

Agents’ choices for consumption, C_t , investment, I_t , and capital, K_t are the outcomes of
 the following maximization problem:

$$\max_{\{C_{t+j}, I_{t+j}, K_{t+j}\}_{j=0}^{\infty}} \sum_{j=0}^{\infty} \frac{C_{t+j}^{1-\gamma} - 1}{1-\gamma} \quad (1)$$

s.t.

$$e^{\varepsilon_{a,t}} K_{t-1}^{\alpha} = C_t + I_t + e^{\varepsilon_{g,t}} \bar{G}, \quad (2)$$

$$K_t = (1 - \delta)K_{t-1} + I_t e^{\varepsilon_{i,t}}. \quad (3)$$

This model contains three exogenous random variables. Using the terminology of Chari et al. (2007), these are an efficiency wedge, $\varepsilon_{a,t}$, an investment wedge, $\varepsilon_{i,t}$, and a government consumption wedge, $\varepsilon_{g,t}$. Consistent with the literature, the variables are generated by the following stochastic process:

$$\varepsilon_{m,t} = \rho_m \varepsilon_{m,t-1} + \sigma_m \eta_{m,t}, m \in \{a, i, g\}, \quad (4)$$

$$\mathbb{E}_t[\eta_{m,t+1}] = 0, \mathbb{E}_t[\eta_{m,t+1}^2] = 1, \text{ and } \mathbb{E}_t[\eta_{m,t+1}\eta_{m^*,t+1}] = 0 \text{ for } m \neq m^*. \quad (5)$$

This economy is represented with the following set of linearized first-order conditions:¹

$$\mathbb{E}_t[c_{t+1} - c_t] = \frac{1 - \beta(1 - \delta)}{\gamma} (\rho_a \varepsilon_{a,t} + (\alpha - 1)k_t) + \frac{1 - \beta(1 - \delta)\rho_i}{\gamma} \varepsilon_{i,t}, \quad (6a)$$

$$\bar{Y}(\varepsilon_{a,t} + \alpha k_{t-1}) = \bar{I} i_t + \bar{C} c_t + \bar{G} \varepsilon_{g,t}, \quad (6b)$$

$$k_t = (1 - \delta)k_{t-1} + \frac{\bar{I}}{\bar{K}} i_t + \frac{\bar{I}}{\bar{K}} \varepsilon_{i,t}, \quad (6c)$$

53 where lower case letters denote variables expressed as a percentage difference from their
54 steady state values and \bar{X} indicates the steady state value of variable X_t .

55 The random disturbances can be interpreted literally as regular exogenous structural
56 disturbances affecting the economy. As illustrated in Chari et al. (2007), however, these
57 wedges can also be seen as manifestations of frictions in more elaborate models or as the part
58 that is not modeled explicitly.² Although they are somewhat general, these three wedges *do*
59 impose restrictions on the model and they differ from each other exactly because of these
60 restrictions. First, none of the wedges appear in all equations, which is typical. Second, the
61 model imposes cross-equation restrictions that depend on the structural parameter values of
62 the model.³

¹Throughout this paper, we focus on linearized systems and treat those as the true data generating process. We do this because most structural empirical macroeconomic models are based on such systems. In principle, one could include ASDs in nonlinear systems as well.

²Moreover, a wedge can be given different interpretations. For example, $\varepsilon_{g,t}$ could be a fixed cost to production or it could be government spending that agents do not value.

³Inoue et al. (2015) provide a formal analysis for using wedges to detect and identify misspecification. Their wedges also only appear in a limited set of equations and, thus, also do impose parameter restrictions.

The three policy functions for this model can be expressed as follows.

$$c_t = A_c(\Psi)k_{t-1} + B_{c,a}(\Psi)\varepsilon_{a,t} + B_{c,i}(\Psi)\varepsilon_{i,t} + B_{c,g}(\Psi)\varepsilon_{g,t}, \quad (7a)$$

$$i_t = A_i(\Psi)k_{t-1} + B_{i,a}(\Psi)\varepsilon_{a,t} + B_{i,i}(\Psi)\varepsilon_{i,t} + B_{i,g}(\Psi)\varepsilon_{g,t}, \quad (7b)$$

$$k_t = A_k(\Psi)k_{t-1} + B_{k,a}(\Psi)\varepsilon_{a,t} + B_{k,i}(\Psi)\varepsilon_{i,t} + B_{k,g}(\Psi)\varepsilon_{g,t}, \quad (7c)$$

63 where Ψ is a vector containing the structural parameters. This system also makes clear that
 64 wedges impose cross-equation restrictions. The $A_j(\Psi)$ and $B_{j,m}(\Psi)$ coefficients are nonlinear
 65 functions of the structural parameters, Ψ .⁴ In linear frameworks, disturbances only differ in
 66 how they affect the economy on impact. After impact they propagate through the economy
 67 in the same way, as described by the A_j s.

68 **Possible misspecification** Misspecification occurs in many different forms. One could
 69 miss a particular disturbance or include one that should not be included. Another possibility
 70 is that a structural disturbance is not incorporated correctly in all model equations. This
 71 is more likely to occur in larger models. However, misspecification is also possible in the
 72 model at hand which has just three equations. For example, the government expenditure
 73 disturbance could very well affect the utility of the agent and/or the production function.
 74 Also, the investment disturbance may affect the depreciation rate.⁵ Another possible mis-
 75 specification is that, contrary to common practice, the structural disturbances are correlated
 76 with each other. Using a New Keynesian business cycle model, Cúrdia and Reis (2012) docu-
 77 ment that structural disturbances are correlated and ignoring this correlation leads to wrong
 78 inference.

79 2.2. Introducing Agnostic Structural Disturbances

ASDs can replace regular structural disturbances or they can be added to the existing set. Adding structural disturbances to model equations is incredibly simple: Each ASD is

⁴See Campbell (1998) for the derivation and discussion of such policy functions.

⁵In Section 4, we provide empirical evidence in support for this possibility.

added to *each* equation with a reduced form coefficient. When we add two ASDs, denoted $\tilde{\varepsilon}_{A,t}$ and $\tilde{\varepsilon}_{B,t}$, to the model of this section, then we get⁶

$$\mathbb{E}_t[c_{t+1} - c_t] = \frac{1 - \beta(1 - \delta)}{\gamma}(\alpha - 1)k_t + [\tilde{\Upsilon}_{1,A}\tilde{\Upsilon}_{1,B}][\tilde{\varepsilon}_{A,t}\tilde{\varepsilon}_{B,t}]', \quad (8a)$$

$$\bar{Y}\alpha k_{t-1} = \bar{I}i_t + \bar{C}c_t + [\tilde{\Upsilon}_{2,A}\tilde{\Upsilon}_{2,B}][\tilde{\varepsilon}_{A,t}\tilde{\varepsilon}_{B,t}]', \quad (8b)$$

$$k_t = (1 - \delta)k_{t-1} + \frac{\bar{I}}{\bar{K}}i_t + [\tilde{\Upsilon}_{3,A}\tilde{\Upsilon}_{3,B}][\tilde{\varepsilon}_{A,t}\tilde{\varepsilon}_{B,t}]', \quad (8c)$$

$$[\tilde{\varepsilon}_{A,t}, \tilde{\varepsilon}_{B,t}]' = \tilde{\varepsilon}_t = P\tilde{\varepsilon}_{t-1} + \tilde{\eta}_t. \quad (8d)$$

80 Each ASD is allowed to enter each equation without any restrictions. Moreover, they enter
 81 the system in a symmetric manner. A priori, there is, thus, no difference between the different
 82 ASDs. It is not restrictive to exclude future realizations of the ASDs from the equations.
 83 What matters is the expectation of these variables and this is captured by the current-period
 84 values as long as the ASDs are first-order Markov processes.

The vector $\tilde{\eta}_t$ contains the ASD innovations. Their standard deviations can be normalized to 1, since the $\tilde{\Upsilon}$ s are reduced-form coefficients. ASD innovations are assumed to be uncorrelated, but the disturbances can be correlated because P does not have to be a diagonal matrix.⁷ Thus, the vector with the ASD innovations, $\tilde{\eta}_t$, satisfies

$$\mathbb{E}_t[\tilde{\eta}_{t+1}] = 0 \text{ and } \mathbb{E}_t[\tilde{\eta}_{t+1}\tilde{\eta}'_{t+1}] = I_2. \quad (9)$$

The policy functions for this model with two ASDs can be expressed as follows.

$$c_t = A_c(\Psi)k_{t-1} + \tilde{B}_{c,A}(\Psi)\tilde{\varepsilon}_{A,t} + \tilde{B}_{c,B}(\Psi)\tilde{\varepsilon}_{B,t}, \quad (10a)$$

$$i_t = A_i(\Psi)k_{t-1} + \tilde{B}_{i,A}(\Psi)\tilde{\varepsilon}_{A,t} + \tilde{B}_{i,B}(\Psi)\tilde{\varepsilon}_{B,t}, \quad (10b)$$

$$k_t = A_k(\Psi)k_{t-1} + \tilde{B}_{k,A}(\Psi)\tilde{\varepsilon}_{A,t} + \tilde{B}_{k,B}(\Psi)\tilde{\varepsilon}_{B,t}, \quad (10c)$$

⁶We have left out the three regular structural disturbances to keep the equations concise.

⁷Correlated innovations can be described by a combination of uncorrelated innovations. Such a setup is fine for agnostic disturbances.

where $A_c(\Psi)$ has the standard solution which does not depend on whether the disturbances are regular or ASDs. The coefficients in these equations are equal to

$$\tilde{B}_{c,m} = \frac{\tilde{\Upsilon}_{1,m} + (\Lambda - A_c(\Psi)) \left(\tilde{\Upsilon}_{3,m} - \frac{\tilde{\Upsilon}_{2,m}}{\bar{K}} \right)}{(\Lambda - A_c(\Psi)) \frac{\bar{C}}{\bar{K}} + \rho - 1}, \quad (11a)$$

$$\tilde{B}_{i,m} = -\frac{\bar{C}\tilde{B}_{c,m} + \tilde{\Upsilon}_{2,m}}{\bar{I}}, \quad (11b)$$

$$\tilde{B}_{k,m} = \frac{\bar{I}}{\bar{K}}\tilde{B}_{i,m} + \tilde{\Upsilon}_{3,m}, \quad (11c)$$

$$\Lambda = \frac{1 - \beta(1 - \delta)}{\gamma}(\alpha - 1). \quad (11d)$$

85 The expressions for the $\tilde{B}_{j,m}(\Psi)$ coefficients illustrate the structural nature of ASDs because
 86 they depend both on the reduced-form $\tilde{\Upsilon}$ coefficients and the structural parameters of the
 87 model, Ψ .

Although the $\tilde{B}_{j,m}(\Psi)$ coefficients depend on Ψ , their values are fully unrestricted. That is, $\tilde{B}_{c,m}(\Psi)$, $\tilde{B}_{i,m}(\Psi)$, and $\tilde{B}_{k,m}(\Psi)$ can take on any set of values by appropriate choice of $\tilde{\Upsilon}_{1,m}(\Psi)$, $\tilde{\Upsilon}_{2,m}(\Psi)$, and $\tilde{\Upsilon}_{3,m}(\Psi)$. Since the $\tilde{B}_{j,m}(\Psi)$ coefficients are unrestricted, an alternative way to implement ASDs is to add them directly to the policy functions with reduced-form coefficients, that is

$$c_t = A_c(\Psi)k_{t-1} + \tilde{B}_{c,A}\tilde{\varepsilon}_{A,t} + \tilde{B}_{c,B}\tilde{\varepsilon}_{B,t}, \quad (12a)$$

$$i_t = A_i(\Psi)k_{t-1} + \tilde{B}_{i,A}\tilde{\varepsilon}_{A,t} + \tilde{B}_{i,B}\tilde{\varepsilon}_{B,t}, \quad (12b)$$

$$k_t = A_k(\Psi)k_{t-1} + \tilde{B}_{k,A}\tilde{\varepsilon}_{A,t} + \tilde{B}_{k,B}\tilde{\varepsilon}_{B,t}, \quad (12c)$$

88 This illustrates that the ASD procedure adds to the policy functions an unobserved compo-
 89 nents block. Describing time-series fully or partly with unobserved components has a rich
 90 history in macroeconomics.⁸ This paper does more than that. Section 4 illustrates how ASDs
 91 can be used as a formal test of the correct specification of regular structural disturbances
 92 and how ASDs can be used to improve upon the specification of structural disturbances.

⁸See, for example, Stock and Watson (1999).

93 The data provide information about the coefficients of the unobserved components block,
94 i.e., the \tilde{B} coefficients. However, to give an economically meaningful interpretation of the
95 ASDs the $\tilde{\Upsilon}$ coefficients are important. The link between the \tilde{B} and the $\tilde{\Upsilon}$ coefficients will
96 be discussed in more detail in Section 2.4.

97 **2.3. ASDs and misspecification**

98 ASDs can detect different types of misspecification. The procedure will indicate an addi-
99 tional structural disturbance is needed if *adding* an ASD improves model fit with the proper
100 adjustment for the additional parameters introduced by the ASD. If *replacing* a regular
101 structural disturbance by an ASD leads to improved model fit (adjusted for the number of
102 parameters), then this indicates that the regular structural disturbance in question either
103 needs to be modified or should not play a role in the empirical model.

104 Cúrdia and Reis (2012) test whether regular structural disturbances are *dynamically* cor-
105 related, that is, the innovations are orthogonal but lagged values of disturbances can affect
106 current values of other disturbances. They find empirical evidence for such dynamic corre-
107 lation.⁹ ASDs can represent the role of correlated disturbances even if P is diagonal, which
108 means that the ASDs are uncorrelated since the innovations of ASDs are assumed to be
109 orthogonal. For example, suppose that both the efficiency wedge, $\varepsilon_{a,t}$ and the investment
110 wedge, $\varepsilon_{i,t}$ are driven by a common component and an idiosyncratic component. Then an
111 empirical model with three ASDs can capture the role of the three different random compo-
112 nents. One vector of $\tilde{\Upsilon}$ coefficients would capture the effect of the common component on
113 the equations which would combine the effects of the $\varepsilon_{a,t}$ and the $\varepsilon_{i,t}$ disturbance. The other
114 two $\tilde{\Upsilon}$ vectors would capture the effects of the idiosyncratic components which would be the
115 separate effect of the wedges.

116 ASDs can also capture measurement error. The ASD system as specified in Equation (8)

⁹Galizia (2015) shows that correlated estimates of the *innovations* of structural disturbances can be a sign of model misspecification. The paper demonstrates that the cross-correlations distorts the estimated variance decomposition of the model and proposes a method to mitigate this problem. As we show below, ASDs also help with getting a lower cross-correlation between the estimated innovations.

117 can capture measurement error in the control variables c_t and i_t .¹⁰ To correctly represent
 118 measurement error in the state variable, k_t one would have to add lagged ASD values to
 119 the system.¹¹ Although ASDs are general enough to encompass measurement error, typical
 120 ASDs differ in a fundamental way from measurement error. In general, ASDs are structural
 121 disturbances and propagate through the system like regular structural disturbances, that is,
 122 according to the $A(\Psi)$ coefficients. Measurement error does not.¹²

123 ASDs are designed to deal with misspecification of structural disturbances that would
 124 distort the $B(\Psi)$ coefficients. Is the ASD procedure also able to deal with misspecification
 125 that affects the $A(\Psi)$ coefficients? Suppose one compares an empirical model with only
 126 one ASD with one that contains one regular structural disturbance and this disturbance
 127 is correctly specified. Moreover, both use $\hat{A}(\Psi)$ which differs from the true $A(\Psi)$. The
 128 ASD specification can still fully represent the correct policy function as long as there is a
 129 $\hat{\Psi}$ such that $\hat{A}(\hat{\Psi}) = A(\Psi)$. The specification with the regular structural disturbance faces
 130 a dilemma. With $\hat{\Psi}$ it gets the A coefficient right, but the B coefficient wrong because it
 131 is improbable that $B(\hat{\Psi}) = B(\Psi)$. If it chooses the correct value for Ψ then it gets B right
 132 but A wrong. The flexibility of the ASD procedure makes it more likely it gets the policy
 133 function coefficients right, not only in terms of the \tilde{B} , but also in terms of the A coefficients.
 134 However, the example shows that this may come at the cost of larger distortions in estimates
 135 of Ψ if the ASD replaces a correctly specified regular structural disturbance.

136 Although our procedure can potentially alleviate misspecification of the $A(\Psi)$ matrix, we
 137 think of our procedure as a first step to understand where the model needs improvement not
 138 as a complete model evaluation.

¹⁰To see this simply replace c_t with $c_{\text{obs},t} + \varepsilon_t$ in Equation (6). After taking expectations one is left with just the current-value of the measurement error term, ε_t .

¹¹Adding lagged values of the ASDs will increase the types of misspecification ASDs can detect. For example, this richer ASD specification could detect whether the period- t value of the productivity disturbance is known in period t , as is commonly assumed, or is known in period $t - 1$, that is, when there are “news” shocks.

¹²When $\tilde{\varepsilon}_{A,t}$ picks up measurement error in c_t or i_t , then the $\tilde{B}_{k,A}$ coefficient associated with $\tilde{\varepsilon}_{A,t}$ would be equal to zero. When an ASD in the enhanced system with lagged ASDs picks up measurement error in k_t , then there is also a set of restrictions such that the $A(\Psi)$ coefficients do not matter for the propagation of $\tilde{\varepsilon}_{A,t}$.

139 2.4. Identification

140 The data provide information about the A and the B coefficients. To understand whether
 141 the A and the B coefficients can be estimated it is useful to think of the model variables as
 142 MA processes. With M ASDs, c_t is a sum of M MA processes. The parameters of each MA
 143 process depend on: $A_c(\Psi)$, $A_k(\Psi)$, $\tilde{B}_{c,m}$, and $\tilde{B}_{k,m}$, $m \in \{1, \dots, M\}$. Thus if one uses just
 144 c_t as an observable, then one can estimate these A and \tilde{B} coefficients, but one would not be
 145 able to estimate the ASD coefficient in the investment policy function, $\tilde{B}_{i,m}$. If one replaces
 146 regular structural disturbances with ASDs, then it may be harder to identify Ψ . That turns
 147 out not to be an issue in our empirical application presented in Section 4.

148 The A s and \tilde{B} s determine the policy functions and moment properties. These may provide
 149 some information on the nature of the ASD. The $\tilde{\Upsilon}$ s indicate how ASD enters each and every
 150 equation. In the empirical application in Section 4, we find that knowing the $\tilde{\Upsilon}$ s is especially
 151 useful for interpreting the different ASDs. So the question arises whether knowing the \tilde{B} s is
 152 enough to determine the $\tilde{\Upsilon}$ s. Equation (11) makes clear that knowing $\tilde{B}_{c,m}$, $\tilde{B}_{i,m}$, and $\tilde{B}_{k,m}$
 153 is necessary but not sufficient to determine $\tilde{\Upsilon}_{1,m}$, $\tilde{\Upsilon}_{2,m}$, and $\tilde{\Upsilon}_{3,m}$.¹³ In addition, one would
 154 need certain combinations of the structural parameters.¹⁴

155 2.5. ASDs versus DSGE-VARs

156 Ireland (2004) and Del Negro et al. (2007) combine a DSGE model with a reduced-form
 157 VAR that contains the observables. There are several key differences between these two
 158 approaches and ours.

159 The DSGE-VAR specification is best compared with the system given in Equation (12)
 160 which adds ASDs to the policy functions. However, an advantage of the ASD procedure is
 161 that one can also obtain the specification given in Equation (8) that determines how the
 162 disturbances affect model equations. This knowledge is helpful in interpreting the nature of

¹³To estimate all three \tilde{B} coefficients one would need data on both consumption and investment.

¹⁴For example, the following expression would need to be identified: $\frac{(\Lambda - A_c(\Psi))}{(\Lambda - A_c(\Psi))^{\frac{C}{K}} + \rho - 1}$.

163 the ASDs as shown in Section 4. This cannot be done with the DSGE-VAR approach.

164 The ASD approach focuses on a particular type of misspecification, which allows it to
165 use aspects of the model that are assumed to be not affected by the misspecification, namely
166 $A(\Psi)$. The DSGE-VAR approach is more ambitious and also directly considers misspeci-
167 fication of $A(\Psi)$. Introducing a VAR into the empirical model means that the number of
168 disturbances necessarily increases by a number equal to the number of variables in the VAR.
169 Moreover, adding a VAR introduces many more parameters unless the number of observables
170 is small. By contrast, our procedure allows for a more parsimonious approach and could con-
171 sist of adding just one new disturbance or replacing one regular structural disturbance with
172 an agnostic structural disturbance.

173 These differences imply that our approach is more efficient in terms of the number of
174 parameters that it has to estimate.¹⁵ The price of parsimony is that our procedure is not
175 designed to detect misspecification unrelated to structural disturbances, that is, misspecifica-
176 tion associated with restrictions imposed by $A(\Psi)$. However, as discussed above the flexibility
177 of our procedure may still alleviate misspecification of $A(\Psi)$. The DSGE-VAR approach ex-
178 plicitly allows misspecification in $A(\Psi)$. However, Chari et al. (2008) point out that the VAR
179 with a finite number of lags that does not contain *all* the model’s state variables is likely
180 to be misspecified. This means that the DSGE-VAR approach cannot deal with all possible
181 misspecifications either.

182 Another difference emerges as the sample size goes to infinity. With the DSGE-VAR
183 approach one has two “competing” empirical specifications, a DSGE model and a VAR.
184 Since every DSGE suffers from at least some minor misspecification, one can expect the VAR
185 to fully take over as the sample size goes to infinity. If that happens, then one is left with
186 a reduced-form model. This will never happen with our approach, since the propagation of
187 state variables will always be determined by $A(\Psi)$.

¹⁵For example, for the popular DSGE model of Smets and Wouters (2007) with 7 observables, a VAR with 4 lags would mean estimating 204 additional coefficients. As discussed in Section 4, the implementation of our procedure for this model means estimating twelve more parameters.

3. Small-Sample Monte Carlo experiments

In this section, we use two small-sample Monte Carlo experiments to demonstrate that ASDs can be used to detect and correct for misspecification in a typical empirical application. As a byproduct, it is shown that the consequences of minor misspecifications in modeling the regular structural disturbances can lead to large distortions in terms of parameter estimates deviating from their true values. Each experiment consists of 1,000 replications. Additional details and results are discussed in Appendix C.

3.1. True model and empirical specifications.

We use the New Keynesian model of Smets and Wouters (2007), the workhorse model of empirical business cycle analysis, to generate the data for each Monte Carlo replication.

The misspecification of the empirical model. The original SW model has seven exogenous random variables. Those are a TFP disturbance, $\varepsilon_{a,t}$, a risk-premium disturbance, $\varepsilon_{b,t}$, a government spending disturbance, $\varepsilon_{g,t}$, an investment-specific disturbance, $\varepsilon_{i,t}$, a monetary policy disturbance, $\varepsilon_{r,t}$, a price mark-up disturbance, $\varepsilon_{p,t}$, and a wage mark-up disturbance, $\varepsilon_{w,t}$. We leave out one of these seven disturbances when generating data for our misspecification experiments. The empirical specification also leaves out one disturbance, *but* not the right one. Every other aspect of the empirical model is correctly specified, including functional forms, specification of the processes for the exogenous random variables, and the values of the parameters that are not estimated.

These are computationally expensive exercises and we only discuss two of the possible forty-two combinations in detail in this section. In the first Monte Carlo experiment, the true *dgp* does not include the monetary policy disturbance, but the empirical model leaves out the investment disturbance instead. In the second disturbance, the empirical model also leaves out the investment disturbance, but it differs from the first in that the true *dgp* does not include the TFP disturbance. In Appendix D, we abstract from small-sampling noise and discuss all forty-two experiments in detail. The appendix also shows that distortions in

214 parameter estimates carry over to implied model properties and explains why we chose these
215 two experiments for this section’s Monte Carlo experiments.

216 **Is this a “minor” misspecification?** When generating the data, we adjust the standard
217 deviation of the disturbance that is incorrectly excluded from the empirical specification to
218 ensure that it is responsible for at most 10% of the volatility for *any* of the six observables
219 used in the estimation. This reduces the quantitative importance of the misspecification.

220 One could argue that a misspecification is only minor if one would not detect it in a
221 typical data set using some model selection criterion such as the marginal likelihood. This is
222 a very strict requirement. Comparing a misspecified model with the true one requires that
223 researchers are aware of the correct specification and test their empirical model against it.
224 Since structural disturbances can enter models in many different ways, researchers may not
225 consider the correct one even if they consider several alternatives. Nevertheless, we implement
226 this test adopting the Bayesian estimation methodology used in Smets and Wouters (2007)
227 with the same priors. Using the marginal data density, the misspecified specification is
228 preferred over the true specification in 17% and 47% of the generated samples for the first
229 and the second Monte Carlo experiment, respectively.

230 **Is this a likely misspecification?** We believe that this type of misspecification is likely
231 to be important in practice even if one includes a large set of structural disturbances. The
232 first reason is that having a large set does not necessarily imply one includes all the true
233 disturbances. Moreover, one does not only need to include all true disturbances, *each* dis-
234 turbance has to enter *each* model equation correctly. For example, a TFP disturbance is
235 typically modeled as a labor-augmenting productivity shock, but productivity changes could
236 affect the production function differently. Moreover, TFP may also affect other aspects of
237 the production process such as the depreciation rate. Moreover, one could argue, that this
238 misspecification is not that likely for the analysis in Smets and Wouters (2007), since SW
239 was preceded by years of empirical analysis by many authors. In Section 4, however, we
240 document that we clearly reject the null that two of the included structural disturbances are

241 correctly specified.

242 **Observables and sample size.** The set of observables used in SW consists of employment,
243 the federal funds rate, the inflation rate, GDP, consumption, investment, and the real wage
244 rate. We exclude the real wage rate so we have the same number of observables as structural
245 disturbances which is consistent with the empirical exercise in SW. We use a sample of typical
246 length, namely 156, which is the same as the number of observations used to estimate the
247 model in Smets and Wouters (2007).

248 **Estimation procedure.** DSGE models are typically estimated with Bayesian techniques,
249 which means that the estimation outcome is a weighted combination of the prior and the
250 empirical likelihood. Misspecification of the empirical model affects the latter. Observed
251 data – and thus misspecification of the likelihood – matter less for posterior estimates with
252 a tight prior. The quality of the estimates will then depend on the quality of the prior. This
253 paper focuses on the question how misspecification affects what the observed data imply for
254 parameter estimates. Thus, we focus on the likelihood and use Maximum Likelihood (ML)
255 estimation. We do impose bounds on the range of parameters considered which alleviates
256 the complexity of the optimization problem.

257 Priors on the standard deviation of structural disturbances typically do not allow for point
258 mass at zero. Ferroni et al. (2015) point out that this biases the results towards a positive
259 role of all structural disturbances. This is not an issue for us, since we use ML estimation. In
260 fact, estimated standard deviations of disturbances that are part of the empirical model but
261 *not* part of the true *dgp* turn out to be often close to zero. Parameter values of the true data
262 generating process are set equal to those of the SW posterior mode. The list of parameters
263 estimated and their interpretation is given in Table 1.

264 **3.2. Evaluating the performance of the ASD approach**

265 In this section, we discuss the results of our Monte Carlo experiments. The outcomes
266 for three different empirical models are compared. The first empirical specification correctly

267 models all regular structural disturbances as in Smets and Wouters (2007). This approach
268 is denoted SW. The second empirical model excludes one regular structural disturbance
269 that is part of and includes one regular structural disturbance that is not part of the true
270 *dgp*. The third empirical model also excludes one of the regular structural disturbances, but
271 replaces it with an ASD. This ASD empirical model is of a more reduced-form nature than
272 the SW specification, but it is not misspecified. That is, there are values of the reduced-form
273 parameters such that it matches the true model.

274 Section 3.2.1 discusses the results when ASDs are used to detect misspecification both
275 when the empirical model is indeed misspecified and when it is not. Section 3.2.2 discusses
276 the ability of ASDs to correct for misspecification.

277 **3.2.1. Using ASDs to detect misspecification**

278 A good test for misspecification has power to reject a misspecified model and rejects a
279 correctly specified model at the chosen significance level. This section documents that ASDs
280 are capable of doing both.

281 **Case I: The empirical model is not correctly specified.** To evaluate whether the ASD
282 procedure can detect misspecification, we use a Likelihood Ratio (LR) test that compares the
283 likelihood of the agnostic empirical specification to the likelihood of the misspecified empirical
284 model. The number of degrees of freedom is equal to ten, since the agnostic specification
285 has ten more parameters.¹⁶ With this procedure, the ASD procedure rejects the misspecified
286 model in all Monte Carlo replications in both experiments. The procedure is, thus, quite
287 powerful in detecting misspecification. The power of the test would decrease if one would
288 use a Bayesian approach, since the common prior would make the posterior of the empirical
289 model with an ASD and the misspecified empirical model more similar and less dependent
290 on the data.

¹⁶We use the formulation of our procedure that adds ASDs directly to the policy functions. This formulation introduces the smallest possible number of additional parameters.

291 **Case II: The empirical model is correctly specified.** For the first Monte Carlo exper-
292 iment, we find that the rejection rate is 21.5% at the 10%-level and 12% at the 5%-level. For
293 the second experiment, these two numbers are 20.9% and 12.6%. The standard error for an
294 estimated fraction is given by $\widehat{f}(1 - \widehat{f})/\sqrt{1000}$, so these differences are significantly different
295 from their theoretical counterpart. Although these small-sample results do not coincide pre-
296 cisely with the theoretical predictions based on large-sample theory, the distortions are not
297 unreasonable. In Appendix C, we document that the histograms of estimated χ^2 statistics
298 are reasonably close to the theoretical (large-sample) χ^2 distribution, but – as indicated by
299 the numbers above – have a slightly fatter upper tail.

300 **3.2.2. Using ASDs to correct for misspecification**

301 The discussion above made clear that the ASD procedure does very well in terms of
302 detecting misspecified models and reasonably well in not rejecting correctly specified models
303 in small samples. In this subsection, we document that the estimates of the structural
304 parameters obtained with the agnostic procedure are much closer to the true values than
305 those obtained with the misspecified empirical model. In fact, they are very similar to those
306 obtained with the correctly specified empirical model with all structural disturbances fully
307 modeled.

308 Table 2 reports the average absolute error of the parameter estimates relative to the true
309 value for the three different empirical models across Monte Carlo experiments. Parameter
310 estimates obtained with the misspecified structural model are substantially worse than those
311 obtained with the correctly specified model. The average of the errors for the misspecified
312 model is more than twice as large as the one for the -fully-specified SW model for several
313 parameters and for both experiments.¹⁷ For the misspecified model, the average errors are
314 typically better for the second than for the first experiment. However, that is not true for all

¹⁷Particular problematic is the standard deviation of the TFP disturbance in the first Monte Carlo experiment for which the average error is almost nine time as large as the one for the correct empirical model. Consistent with the broader investigation of Appendix D, this disturbance often takes over the role of the wrongly excluded structural disturbance.

315 parameters. For example, the average error for σ_c is substantially higher in the second exper-
316 iment, whereas there is only a modest increase for the correctly specified model. Appendix
317 D, which discusses the consequences of misspecification for all forty-two experiments, shows
318 that the substantial distortions in parameter estimates reported here are not atypical and
319 also distort implied model properties such as business cycle moments and IRFs.

320 For the first Monte Carlo experiment, the average errors for the agnostic setup and the
321 SW specification are very similar. Although only slightly, the average error is actually lower
322 for the agnostic specification for ten of the twenty-seven parameters. Note that the agnostic
323 specification is not misspecified, but has a disadvantage relative to the SW specification
324 since it uses a reduced-form approach and contains ten more parameters. Nevertheless, the
325 efficiency loss turns out to be very minor.

326 For the second Monte Carlo experiment, the SW specification comes with some noticeable
327 efficiency advantages for some parameter estimates. Nevertheless, estimates obtained with
328 the ASD procedure are still much better than the one obtained with the misspecified model.

329 Figures 1 and 2 plot histograms characterizing the distribution of the parameter esti-
330 mates across Monte Carlo replications for a selected set of parameters. Each panel reports
331 the results for the fully-specified SW model (dark line and dots), the agnostic procedure
332 (white bars), and the misspecified model (blue/dark bars). The figures document that the
333 distributions of estimates obtained with the SW specification and the agnostic procedure are
334 both qualitatively and quantitatively very similar. By contrast, the distribution of estimates
335 obtained with the misspecified empirical model can be vastly different. For example, Panel
336 a of Figure 1 documents that the distribution of estimates of the capital share parameter,
337 α , displays a strong downward bias when the misspecified empirical model is used. The
338 associated mean is equal to 0.09, whereas the true value is equal to 0.19. The figure also
339 documents that a large number of estimates are clustered at the imposed lower bound. That
340 is, by imposing bounds we limited the distortions due to misspecification. For α , the leftward
341 shift is so large, that there is little overlap between the distribution of the estimates based on
342 the misspecified model and the other two empirical models. Bunching at the lower or upper
343 bound is more pervasive for the first experiment, but also observed for the second.

344 For the parameters considered in these figures, the distribution of estimates for the ag-
345 nostic and the fully-specified SW specification are almost always centered around the true
346 parameter value. In principle, there could be a small sample bias, since this is a complex
347 nonlinear estimation problem. The full set of results, discussed in Appendix C, do indeed
348 indicate that there is a bias for some parameters. In those cases, the bias is similar for the
349 estimator based on the fully-specified specification and the agnostic one. An example of a
350 parameter that is estimated with bias is the labor supply elasticity with respect to the real
351 wage, σ_l . Its true value is equal to 1.92. In the first experiment, the average estimate across
352 the Monte Carlo replications is equal to 1.84 for the SW and 1.71 for the agnostic specifica-
353 tion. By contrast, the associated average estimate is equal to 0.27 for the misspecified model,
354 which indicates a bias of a much larger magnitude.

355 **4. Are the SW disturbances the right ones for US data?**

356 In this section, we first apply the ASD procedure to test the restrictions imposed by the
357 SW structural disturbances with the US postwar data used by SW. We document that the
358 restrictions imposed by the risk premium and the investment-specific technology disturbance
359 are rejected by the ASD procedure. Next, we use model selection procedures to determine
360 the number of ASDs to include and to construct a more concise specification that excludes
361 the agnostic disturbances from some model equations. To conclude, we interpret the nature
362 of the agnostic structural disturbances by examining the sign and magnitude of their asso-
363 ciated coefficients in model equations and their IRFs. Appendix E provides more detailed
364 information on our empirical analysis and additional results.

365 **4.1. Testing the Smets-Wouters disturbance restrictions**

366 Since SW use a Bayesian estimation procedure, we do the same. Implementing the ASD
367 procedure only requires a minor modification of the Dynare program that estimates the model
368 for the original SW specification. Replacing an SW regular structural disturbance with an
369 ASD introduces a 13×1 $\tilde{\Upsilon}$ vector but only twelve additional parameters to estimate, since

370 the standard deviation of the ASD innovation is normalized to 1.

371 Suppose an original SW disturbance enters the j^{th} equation with coefficient $\tilde{\Upsilon}_j(\Psi)$. When
372 it is replaced with an ASD, then we set the prior for the ASD coefficient in the j^{th} equation to
373 a Normal with a mean equal to $\tilde{\Upsilon}_j(\Psi)$ with Ψ evaluated at SW prior means. By centering the
374 priors of the agnostic coefficients around the SW restrictions, we favor the SW specification.
375 However, the means of these priors hardly matter and our results are robust to setting the
376 prior mean equal to zero for all coefficients. The standard deviations of the prior distributions
377 for the $\tilde{\Upsilon}$ coefficients are set equal to 0.5. This implies very uninformed priors, since the model
378 is linear in log variables. As a robustness check we also consider a standard deviation equal
379 to 0.1 and we find very similar results.

380 The specification that replaces a regular disturbance with an agnostic one encompasses
381 the original specification which gives it an advantage in terms of achieving a better fit. The
382 additional parameters, however, act as a penalty term in the marginal data density. Table 3
383 reports the marginal data densities for the original SW specification and for specifications
384 in which the indicated regular structural disturbances is replaced by an ASD. Overall, these
385 outcomes are quite supportive of the original SW specification as the SW restrictions are
386 preferred for five of the seven structural disturbances. But the results for the risk-premium
387 and the investment specific disturbance indicate that improvement is possible.

388 4.2. Obtaining our preferred model with ASDs

389 These results do not necessarily imply that we should exclude the structural risk-premium
390 and investment disturbance. After all, it is possible that a model that includes agnostic dis-
391 turbances *as well as* these two SW structural disturbances has an even higher marginal data
392 density. Moreover, ASDs add quite a few extra parameters which may make interpretation
393 more difficult. The next step of the ASD procedure is to use a model selection procedure.

394 There are different model selection procedures one can use to obtain a preferred specifi-
395 cation. Our procedure is described in detail in Appendix E.2. The chosen model is one that
396 excludes the SW risk premium as well as the SW investment disturbance, it includes three

397 ASDs, and imposes several zero restrictions on the ASD coefficients.

398 **4.3. Giving the ASDs an economic interpretation**

399 ASDs are agnostic by nature. The model selection procedure also does not use any eco-
400 nomic reasoning. Here we will show how the estimation results, such as parameter estimates
401 of ASD coefficients and IRFs, can be used to give a meaningful interpretation to the ASDs.
402 We will argue that one of the three selected ASDs can be interpreted as an investment-specific
403 disturbance, but with some quite striking differences from the regular one used in the lit-
404 erature and in SW. We will refer to this ASD as the agnostic “investment-modernization
405 disturbance.” The second ASD has features in common with the SW risk-premium distur-
406 bance and with a preference disturbance, but is different from both. We will refer to this
407 ASD as the agnostic “Euler disturbance.” The role of the third ASD is quantitatively less
408 important than the other two. It mainly affects wage growth and is associated with a more
409 efficient use of capital. We will refer to this ASD as the “capital-efficiency wage mark-up
410 disturbance.” By assigning names to agnostic disturbances, we may open ourselves to criti-
411 cism. Our main reason for assigning these names is that we want to make clear that agnostic
412 disturbances are in principle theory-free, *and yet* allow the researcher to go one step further,
413 towards giving an economic interpretation to them.

414 **4.3.1. The agnostic investment-modernization disturbance, $\tilde{\varepsilon}_{B,t}$**

415 In the SW model, the investment-specific technology disturbance shows up in the in-
416 vestment Euler equation and in the capital accumulation equation. One of our agnostic
417 disturbances, $\tilde{\varepsilon}_{B,t}$, also shows up in these two equations.¹⁸ The only other equation in which
418 $\tilde{\varepsilon}_{B,t}$ appears is the equation that relates capacity utilization to the rental rate of capital.
419 These findings indicate that $\tilde{\varepsilon}_{B,t}$ could be interpreted as an investment-specific productivity

¹⁸In our computer programs, the ASDs are referred to as agnA, agnB, and agnC. The interpretation for agnB is the most straightforward so we discuss this one first. We could have relabeled it as $\tilde{\varepsilon}_{A,t}$, but chose not to do so to emphasize that labels for ASDs are arbitrary.

420 disturbance. Furthermore, as documented in Table 4, $\tilde{\varepsilon}_{B,t}$, plays an important role for the
421 volatility of investment. Specifically, it explains 70% of the volatility of investment growth
422 compared to 82.1% for the investment-specific disturbance in the SW model. Interestingly,
423 $\tilde{\varepsilon}_{B,t}$ is not important for the volatility of capital. Specifically it only explains 2.37% of the
424 volatility of the capital stock, whereas the SW investment disturbance explains 32.5%. Thus,
425 if $\tilde{\varepsilon}_{B,t}$ is an investment-specific disturbance, then it is not a typical one.

426 Figure 3 plots the IRFs of our agnostic disturbance and the SW investment-specific dis-
427 turbance. This graph documents there are some remarkable differences. The SW investment
428 disturbance generates a typical business cycle with key aggregates moving in the same direc-
429 tion. A positive agnostic investment disturbance also leads to a strong comovement between
430 output and investment, but leads to a *reduction* in consumption and capital.¹⁹ Also, whereas
431 capacity utilization decreases in the SW model, our specification indicates an increase.

432 To understand these differences and to explain why we still think that $\tilde{\varepsilon}_{B,t}$ is an investment-
433 specific disturbance, we have to take a closer look at the relevant equations and how $\tilde{\varepsilon}_{B,t}$
434 affects these equations differently than the SW investment specific disturbance, $\varepsilon_{i,t}$. The
435 three relevant equations are the following:²⁰

Smets-Wouters investment-specific disturbance, $\varepsilon_{i,t}$

$$\text{Investment Euler: } i_t = i_1(\Psi) i_{t-1} + (1 - i_1(\Psi)) \mathbb{E}_t [i_{t+1}] + \varepsilon_{i,t}, \quad (13)$$

$$\text{Utilization: } z_t = z_1(\Psi) r_t^k, \quad (14)$$

$$\text{Capital: } k_t = k_1(\Psi) k_{t-1} + (1 - k_1(\Psi)) i_t + \mathbf{k}_2(\Psi) \varepsilon_{i,t}, \quad \mathbf{k}_2(\Psi) > \mathbf{0}. \quad (15)$$

¹⁹Justiano et al. (2010) also report a negative consumption response to an investment disturbance, but only for the first five periods. As discussed in Ascari et al. (2016), most models would predict a countercyclical consumption response to an investment disturbance. The SW model overturns this property due to a sufficiently high degree of price and wage stickiness. Our agnostic approach implies similar estimates for price and wage stickiness, but still indicates that the data prefer a countercyclical consumption response.

²⁰The subscripts of the coefficients of the agnostic disturbance refer to the SW equation number. For example, $\tilde{\Upsilon}_{3,B} \tilde{\varepsilon}_{B,t}$ is the term added to Equation (3) of SW. i_t is the investment level, r_t^k the rental rate of capital, z_t the utilization rate, $\varepsilon_{i,t}$ the SW investment-specific investment disturbance, and Ψ is the vector with structural parameters.

Agnostic investment-modernization disturbance, $\tilde{\varepsilon}_{B,t}$

$$\text{Investment Euler: } i_t = i_1(\Psi) i_{t-1} + (1 - i_1(\Psi)) \mathbb{E}_t [i_{t+1}] + \tilde{\Upsilon}_{3,B} \tilde{\varepsilon}_{B,t}, \tilde{\Upsilon}_{3,B} > 0, \quad (16)$$

$$\text{Utilization: } z_t = z_1(\Psi) r_t^k + \tilde{\Upsilon}_{7,B} \tilde{\varepsilon}_{B,t}, \tilde{\Upsilon}_{7,B} < 0, \quad (17)$$

$$\text{Capital: } k_t = k_1(\Psi) k_{t-1} + (1 - k_1(\Psi)) i_t + \tilde{\Upsilon}_{8,B} \tilde{\varepsilon}_{B,t}, \tilde{\Upsilon}_{8,B} < 0. \quad (18)$$

436 The reason for the striking differences between the IRFs of our ASD and the SW in-
 437 vestment disturbance is that our unrestricted approach lets the agnostic investment-specific
 438 disturbance appear in the capital accumulation equation *without* restrictions. That is, the
 439 sign of the coefficient of $\tilde{\varepsilon}_{B,t}$, $\tilde{\Upsilon}_{8,B}$, is unrestricted, but the coefficient of $\varepsilon_{i,t}$ in the SW spec-
 440 ification, $k_2(\Psi)$ is restricted by the values of the structural parameters, Ψ . The outcome is
 441 that the posterior mean of $\tilde{\Upsilon}_{8,B}$ has the *opposite* sign relative to $k_2(\Psi)$ and the 90% HPD
 442 does not include 0.

443 This means that a reduction in the cost of transforming current investment into capital
 444 goes together with increased depreciation of the existing capital stock in our specification.
 445 In the SW model, an investment-specific disturbance does not affect the economic viability
 446 of the existing capital stock. Our agnostic approach questions this assumption and suggests
 447 that the investment-specific productivity disturbance goes together with scrapping of older
 448 vintages. This is the reason why we refer to it as an agnostic investment-modernization
 449 disturbance.

450 In the SW model, capacity utilization is proportional to the rental rate and there are
 451 no shocks that can affect this relationship. An accelerated depreciation of the capital stock
 452 increases the rental rate, which in turn would induce an increase in the utilization rate. In
 453 our agnostic specification, this relationship is dampened somewhat, since a positive agnostic
 454 disturbance has a *direct* negative impact on capacity utilization, since it enters the capacity
 455 utilization with a negative coefficient. The overall effect is still an increase in capacity uti-
 456 lization. It seems plausible that scrapping of old vintages goes together with higher utilization
 457 of the remaining capital stock.

458 **4.3.2. The agnostic Euler disturbance, $\tilde{\varepsilon}_{A,t}$**

459 The agnostic disturbance $\tilde{\varepsilon}_{A,t}$ appears in eight equations. The key equation is the Euler
460 equation for bonds, because excluding the disturbance from this equation leads to by far the
461 largest drop in the marginal data density. This suggests that it could have key characteristics
462 in common with a preference or a risk-premium disturbance. This view is also supported by
463 Table 4 which documents that $\tilde{\varepsilon}_{A,t}$ is important for the same variables as the SW risk-premium
464 disturbance. However, this agnostic disturbance also has some quite different characteristics
465 from both. Therefore, we will adopt an alternative name and refer to it as the agnostic Euler
466 disturbance. For the interpretation of $\tilde{\varepsilon}_{A,t}$, it is important to understand the differences in
467 impact of a regular preference and a regular (bond) risk-premium disturbance.

468 **Difference between a preference and (bond) risk-premium disturbance.** Smets
469 and Wouters (2003) include a preference disturbance which affects current utility. This
470 means it affects the marginal rate of substitution and, thus, *all* Euler equations. By contrast,
471 Smets and Wouters (2007) include instead a (bond) risk premium that introduces a wedge
472 between the policy rate and the required rate of return on bonds without affecting other Euler
473 equations.²¹ Both disturbances have a strong impact on current consumption when prices
474 are sticky. A positive preference disturbance reduces the attractiveness of *all* types of saving
475 including investment. A positive risk-premium disturbance only makes savings in bonds less
476 attractive. That is, it induces a desire to substitute out of bonds and into investment, in
477 addition to an increase in consumption. Thus, a preference disturbance leads to a negative
478 comovement of consumption and investment, whereas a (bond) risk-premium disturbance
479 leads to a positive comovement.

480 Also, a preference disturbance affects output in both the flexible-price and the sticky-price
481 part of the model, whereas a risk-premium disturbance has no affect on key aggregates such
482 as consumption and output in the flexible price part of the SW model.

²¹If a preference disturbance is added to the specification of Smets and Wouters (2007), then the marginal data density drops from -922.40 to -923.57 and the preference disturbance plays virtually no role.

483 **Is $\tilde{\varepsilon}_{A,t}$ a preference, a risk-premium, or another type of disturbance?** Figure 4 plots
484 the IRFs of the SW risk-premium and our agnostic disturbance. The figure documents that
485 both generate a regular business cycle with positive comovement for output, consumption,
486 investment, and hours. The positive comovement suggest that the agnostic disturbance is
487 a bond risk-premium disturbance and not a preference disturbance. However, the agnostic
488 disturbance has a strong impact on flexible-price output which is inconsistent with it being
489 a (bond) risk-premium disturbance and consistent with it being a preference disturbance.
490 Since this ASD differs from both a preference and a risk-premium disturbance, we come up
491 with a new term, namely the Euler disturbance.

492 To better understand the nature of the agnostic Euler disturbance, we take a closer look
493 at the equations in which $\tilde{\varepsilon}_{A,t}$ enters. It appears in the aggregate budget constraint, the bond
494 Euler equation, the investment Euler equation, the capital value equation, the utilization rate
495 equation, the price mark-up equation, the rental rate of capital equation, and the Taylor rule.
496 Although $\tilde{\varepsilon}_{A,t}$ affects quite a few different aspects of the model, the interpretation is eased
497 by the fact that its role is minor in most of the eight equations in the sense that allowing
498 it to enter these equations only has a minor quantitative impact on the behavior of model
499 variables or only affects the qualitative behavior of one or two variables without affecting the
500 behavior of the key macroeconomic variables.

501 Specifically, to understand the role of $\tilde{\varepsilon}_{A,t}$ on key macroeconomic aggregates we can re-
502 strict ourselves to the Taylor rule and the three model equations that are relevant for the
503 savings/investment decisions, which are the bond Euler equation, the investment Euler equa-
504 tion, and the capital value equation. As in SW, we use the bond Euler equation to substitute
505 the marginal rate of substitution out of the capital valuation equation. While the SW bond
506 risk-premium disturbance, $\varepsilon_{b,t}$, does *not* appear in the original capital valuation equation, it
507 does show up *after* this substitution has taken place. Moreover, it appears in these two equa-
508 tions with the exact same coefficient as the nominal interest rate for bonds, r_t . By contrast,
509 *after* substituting out the marginal rate of substitution in the capital value equation, a pref-
510 erence disturbance would *no longer* appear in the capital valuation equation. The following
511 set of equations documents how the SW risk-premium and our agnostic Euler disturbance

512 enter these equations:²²

Smets-Wouters risk premium, ε_t^b

$$\begin{aligned} \text{Bond Euler:} \quad c_t &= c_1(\Psi) c_{t-1} + (1 - c_1(\Psi)) \mathbb{E}_t [c_{t+1}] + c_2(\Psi) (l_t - \mathbb{E}_t [l_{t+1}]) \\ &\quad - c_3(\Psi) (\mathbf{r}_t - \mathbf{E}_t [\pi_{t+1}] + \varepsilon_{b,t}), c_3(\Psi) > 0, \end{aligned} \quad (19)$$

$$\text{Inv. Euler:} \quad i_t = i_1(\Psi) i_{t-1} + (1 - i_1(\Psi)) \mathbb{E}_t [i_{t+1}] + \varepsilon_{i,t}, \quad (20)$$

$$\begin{aligned} \text{Valuation:} \quad q_t &= q_1 \mathbb{E}_t [q_{t+1}] + (1 - q_1) \mathbb{E}_t [r_{t+1}^k] \\ &\quad - (\mathbf{r}_t - \mathbf{E}_t [\pi_{t+1}] + \varepsilon_{b,t}), \end{aligned} \quad (21)$$

$$\begin{aligned} \text{Policy rate:} \quad r_t &= \rho r_{t-1} + (1 - \rho) \{r_\pi + r_Y (y_t - y_t^p)\} \\ &\quad + r_{\Delta y} [(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)] + \varepsilon_{r,t}. \end{aligned} \quad (22)$$

Agnostic Euler disturbance, $\tilde{\varepsilon}_{A,t}$

$$\begin{aligned} \text{Bond Euler:} \quad c_t &= c_1(\Psi) c_{t-1} + (1 - c_1(\Psi)) \mathbb{E}_t [c_{t+1}] + c_2(\Psi) (l_t - \mathbb{E}_t [l_{t+1}]) \\ &\quad - c_3(\Psi) (\mathbf{r}_t - \mathbf{E}_t [\pi_{t+1}]) - \tilde{\Upsilon}_{2,A} \tilde{\varepsilon}_{A,t}, \tilde{\Upsilon}_{2,A} > 0, \end{aligned} \quad (23)$$

$$\text{Inv. Euler:} \quad i_t = i_1(\Psi) i_{t-1} + (1 - i_1(\Psi)) \mathbb{E}_t [i_{t+1}] + \varepsilon_{i,t} - \tilde{\Upsilon}_{3,A} \tilde{\varepsilon}_{A,t}, \tilde{\Upsilon}_{3,A} > 0, \quad (24)$$

$$\begin{aligned} \text{Valuation:} \quad q_t &= q_1 \mathbb{E}_t [q_{t+1}] + (1 - q_1) \mathbb{E}_t [r_{t+1}^k] \\ &\quad - (\mathbf{r}_t - \mathbf{E}_t [\pi_{t+1}]) - \tilde{\Upsilon}_{4,A} \tilde{\varepsilon}_{A,t}, \tilde{\Upsilon}_{4,A} > 0, \end{aligned} \quad (25)$$

$$\begin{aligned} \text{Policy rate:} \quad r_t &= \rho r_{t-1} + (1 - \rho) \{r_\pi + r_Y (y_t - y_t^p)\} \\ &\quad + r_{\Delta y} [(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)] + \varepsilon_{r,t} + \tilde{\Upsilon}_{14,A} \tilde{\varepsilon}_{A,t}, \tilde{\Upsilon}_{14,A} > 0. \end{aligned} \quad (26)$$

513 Our ASD appears in the bond Euler equation and the capital valuation equation and it
514 shows up with the same sign as the SW risk-premium disturbance. This supports the view
515 that our ASD is similar to a risk-premium disturbance. Nevertheless, one could argue that

²²In these equations, c_t is consumption, l_t is hours worked, r_t is the nominal policy rate, π_t is the inflation rate, q_t is the price of capital, y_t is output, and y_t^p is output in the flexible-price economy. Also see the information given in footnote 20.

516 the ASD is a preference and not a bond risk-premium disturbance for the following reasons.
517 Although $\tilde{\Upsilon}_{4,A}$ has the right sign for a risk-premium coefficient, its magnitude, evaluated
518 using the posterior mean, is way too small.²³ The 90% HPD interval of the coefficient of $\tilde{\varepsilon}_{A,t}$
519 in the capital valuation equation, $\tilde{\Upsilon}_{4,A}$, includes zero and setting the coefficient equal to zero
520 has very little impact on model properties and virtually none on the marginal data density.

521 As pointed out above, a preference disturbance generates consumption and investment
522 responses that move in opposite directions. Our ASD predicts responses in the same direction
523 even if we impose that the ASD does not enter the capital valuation equation (after substi-
524 tuting out the MRS). The reason for the positive comovement is that our ASD also enters
525 the investment Euler equation. The investment Euler equation is a dynamic equation, but
526 its dynamic aspects are due solely to investment adjustment costs.²⁴ Our agnostic approach
527 indicates that the structural disturbance that plays a key role in the bond Euler equation
528 should also appear in the investment Euler equation. In fact, it is the first equation chosen
529 in our specific-to-general model selection procedure.

530 What could this agnostic disturbance represent? A simple explanation is that it is a
531 preference disturbance that is correlated with an investment-specific disturbance. This dis-
532 turbance appears directly in the Taylor rule with a negative coefficient. This means that the
533 central bank responds more aggressively to business cycle fluctuations induced by this Euler
534 disturbance. Without this effect on the Taylor rule this disturbance would have a stronger
535 impact on economic aggregates and inflation would no longer be procyclical.

536 4.3.3. The agnostic capital-efficiency wage mark-up disturbance, $\tilde{\varepsilon}_{C,t}$

537 The third ASD chosen by our model selection criterion increases the total number of
538 structural disturbances to eight, that is, one more than the number in the SW specification.
539 Thus, this ASD cannot be interpreted as a replacement of a SW disturbance.

²³If our ASD is a risk-premium disturbance, then $\tilde{\Upsilon}_{4,A}/\tilde{\Upsilon}_{2,A}$ should be equal to $1/c_3(\Psi)$, but using posterior means, we find that $\tilde{\Upsilon}_{4,A}/\tilde{\Upsilon}_{2,A} = 3.3$, whereas $1/c_3(\Psi) = 7.27$, substantially higher.

²⁴Adjustment costs are zero in the steady state, which implies that neither a preference disturbance nor a risk-premium disturbance appear in a *linearized* investment Euler equation. A preference disturbance would appear in the original *nonlinear* equation.

540 This third ASD, $\tilde{\varepsilon}_{C,t}$, appears in five equations and the most important one (in terms
541 of impact on the marginal data density) is the wage-adjustment equation. It also shows
542 up into three equations related to capital, namely the capital accumulation equation, the
543 capital utilization equation, and the capital-valuation equation. Finally, it appears in the
544 economy-wide budget constraint, although the impact on the latter is minor.

545 Given its impact on the wage equation, this ASD could very well also be a wage distur-
546 bance. Figure 5 plots its IRFs for $\tilde{\varepsilon}_{C,t}$ and for $\varepsilon_{w,t}$ in the SW and in our specification with
547 three ASDs and the SW risk premium and investment disturbance excluded. The IRFs of
548 $\varepsilon_{b,t}$ in the two specifications generate a similar business cycle, also quantitatively. A positive
549 shock to $\tilde{\varepsilon}_{C,t}$ also induces a recession with a reduction in output, investment, and employ-
550 ment. However, it leads to an increase in installed capital and capital services although the
551 latter less than the first. In contrast to the SW $\varepsilon_{w,t}$ shock it goes together with a decrease
552 in the price of capital.

553 $\tilde{\varepsilon}_{C,t}$ is an AR(1) process, and the posterior mean of the auto-regressive coefficient is equal
554 to 0.19. The SW $\varepsilon_{w,t}$ disturbance is a very persistent ARMA(1,1) process. The presence of
555 $\tilde{\varepsilon}_{C,t}$ in the empirical model strongly reduces the coefficient of the MA component of $\varepsilon_{w,t}$, but
556 has little impact on the AR component.²⁵

557 Including $\tilde{\varepsilon}_{C,t}$ in the empirical specification does not reduce the role of $\varepsilon_{w,t}$ for fluctuations
558 of key variables. $\varepsilon_{w,t}$ remains the most important disturbance for key economic aggregates.
559 The only exception is the wage *growth* rate. In the SW specification $\varepsilon_{w,t}$ explains 61.6% of
560 the volatility of wage growth, whereas it only explains 13.3% in our preferred specification.
561 This role is clearly taken over by $\tilde{\varepsilon}_{C,t}$ which explains 53.5% of wage growth volatility. $\tilde{\varepsilon}_{C,t}$
562 also plays a nontrivial role for fluctuations in the capital stock, capacity utilization, and the
563 rental rate of capital, explaining 9.8%, 14.7%, and 13.1%, of total variability respectively.

564 The results indicate that this agnostic disturbance increases the wage mark-up and is
565 associated with a lower price of capital and an increased (use of the) capital stock. One

²⁵Specifically, with $\tilde{\varepsilon}_{C,t}$ included in the empirical specification the posterior means of the AR and the MA coefficients of $\varepsilon_{w,t}$ are equal to 0.97 and 0.59, respectively. Estimates with the SW specification for these two numbers are 0.97 and 0.85.

566 possible explanation is that the increase in the level of used capital (possibly induced by
567 lower prices) comes at the cost of higher wage rates. That is, to operate this larger capital
568 stock, firms have to pay a higher wage rate, perhaps in terms of an overtime premium. The
569 relevant equations are the following:²⁶

Agnostic capital-efficiency-wage-mark-up disturbance, $\tilde{\varepsilon}_{C,t}$

$$\begin{aligned} \text{Valuation:} \quad q_t &= q_1 \mathbb{E}_t [q_{t+1}] + (1 - q_1) \mathbb{E}_t [r_{t+1}^k] \\ &\quad - (\mathbf{r}_t - \mathbf{E}_t [\pi_{t+1}]) - \tilde{\Upsilon}_{4,C} \tilde{\varepsilon}_{C,t}, \quad \tilde{\Upsilon}_{4,C} < \mathbf{0}, \end{aligned} \quad (27)$$

$$\text{Utilization:} \quad z_t = z_1 (\Psi) r_t^k + \tilde{\Upsilon}_{7,C} \tilde{\varepsilon}_{C,t}, \quad \tilde{\Upsilon}_{7,C} > \mathbf{0}, \quad (28)$$

$$\text{Capital:} \quad k_t = k_1 (\Psi) k_{t-1} + (1 - k_1 (\Psi)) i_t + \tilde{\Upsilon}_{8,C} \tilde{\varepsilon}_{C,t}, \quad \tilde{\Upsilon}_{8,C} > \mathbf{0}, \quad (29)$$

$$\begin{aligned} \text{Wage:} \quad w_t &= w_1 w_{t-1} + (1 - w_1) (\mathbb{E}_t [w_{t+1} + \pi_{t+1}]) \\ &\quad - w_2 \pi_t + w_3 \pi_{t-1} - w_4 \mu_{w,t} + \tilde{\Upsilon}_{13,C} \tilde{\varepsilon}_{C,t}, \quad \tilde{\Upsilon}_{8,C} > \mathbf{0}. \end{aligned} \quad (30)$$

570 4.4. Correlation estimated innovations

571 Estimated innovations are supposed to be orthogonal to each other and display no auto-
572 correlation. In practice this is often not the case. As shown in Appendix E.3, the ASD system
573 does a substantially better job than the SW system regarding cross-correlations. Both $\tilde{\eta}_{A,t}$
574 and $\tilde{\eta}_{B,t}$ are less correlated with other innovations than their SW counterparts $\eta_{b,t}$ and $\eta_{i,t}$.
575 Moreover, the cross-correlations of the regular structural disturbances that are present in
576 both specifications are also less correlated. Specifically, whereas the SW has nine correlation
577 coefficients that are significantly different from zero at the 10% level for its seven innovations,
578 the ASD system has four significant correlation coefficients for its eight innovations and only
579 two if we exclude the eighth innovation of the ASD system that is associated with $\tilde{\varepsilon}_{C,t}$.

580 The ASD specification also does better regarding the auto-correlation of the innovations.

²⁶We leave out the overall budget constraint since the role of the disturbance in this equation is very minor, but its impact in this equation is like a contractionary fiscal expenditure shock. w_t is the real wage rate and $\mu_{w,t}$ is the real wage mark-up, i.e., the difference between the wage rate and the marginal rate of substitution between consumption and leisure. Also see footnote 20 for additional information.

581 In the SW system, four of the estimated innovations display significant auto-correlation at
582 the 10% level. That number reduces to two in the ASD system and one of the significant
583 coefficients is for the innovation associated with the additional disturbance, $\tilde{\varepsilon}_{C,t}$, for which
584 there is no counterpart in the SW system.

585 **5. Concluding comments**

586 Structural disturbances play a key role in modern business cycles models. Thus, it is
587 important to introduce them correctly. Having wrong formulations will lead to the wrong
588 inference on what type of disturbances matter most for the fluctuations of key economic
589 variables. One of the main objectives of structural models is to do policy analysis. Deriving
590 optimal fiscal and monetary policy correctly also depends crucially on formulating structural
591 disturbances correctly since these are important ingredients of optimal policy rules.

592 This paper shows that misspecifications can also lead to substantial distortions in param-
593 eter estimates and implied model properties. Obviously, the analysis of government policies
594 will be flawed if parameter estimates are incorrect. For example, the impact of monetary pol-
595 icy on economic aggregates in New Keynesian models depend crucially on getting parameters
596 related to the degree of price and wage stickiness right.

597 The development of MCMC techniques has made it possible to estimate larger models
598 with a larger set of observables. To avoid singularity issues this also requires including more
599 disturbances which enhances the challenge to model them all correctly. ASDs can help.
600 First, they can be used to test whether the specification of a regular structural disturbance
601 is correct and if found problematic can provide insights on how to improve its specification.
602 Researchers can also simply add ASDs to the set of structural disturbances without having
603 any concern about these introducing misspecification.

604 Focusing on the misspecification of disturbances is only a first step in a proper evaluation
605 of a structural model. Moreover, economists are often more interested in how the model itself
606 magnifies and propagates shocks than in what created the initial disruption. Our procedure
607 is helpful in this regard. By being more agnostic about the nature of structural disturbances

608 one is less likely to distort the analysis of what one is ultimately interested in.

Table 1: Parameter explanations

| | |
|----------------|--|
| α | Capital share |
| σ_c | Inverse IES of consumption |
| Φ | Fixed cost in production |
| ϕ | Elasticity of adjustment cost function |
| λ | Degree of consumption habits |
| ξ_w | Degree of wage rigidity |
| σ_ℓ | Inverse IES of leisure |
| ξ_p | Degree of price rigidity |
| ι_w | Degree of indexation for wages |
| ι_p | Degree of indexation for prices |
| ψ | Elasticity of capital utilization adj. cost function |
| r_π | Taylor rule coefficient on inflation |
| ρ | Degree of interest rate smoothing in Taylor rule |
| r_y | Taylor rule coefficient on output gap |
| $r_{\Delta y}$ | Taylor rule coefficient on change in output gap |

| | |
|------------|---|
| | for $j \in \{a, b, g, i, r, p, w\}$: |
| ρ_j | Persistence of exogenous disturbance j |
| σ_j | Standard deviation of exogenous disturbance j |
| | for $j \in \{p, w\}$: |
| μ_j | MA coefficient of exogenous disturbance j |

Notes. The table reports the parameters of the SW model that are estimated and their interpretation. The list of exogenous disturbances is given in the text.

Table 2: Average absolute errors across Monte Carlo experiments

| true value | | average error first MC | | | average error second MC | | |
|----------------|------|------------------------|----------|-------|-------------------------|----------|-------|
| | | misspecified | agnostic | SW | misspecified | agnostic | SW |
| α | 0.19 | 0.098 | 0.035 | 0.028 | 0.056 | 0.048 | 0.037 |
| σ_c | 1.39 | 0.384 | 0.246 | 0.191 | 0.540 | 0.288 | 0.226 |
| Φ | 1.61 | 0.217 | 0.212 | 0.191 | 0.192 | 0.212 | 0.164 |
| ϕ | 5.48 | 1.793 | 1.326 | 0.899 | 1.429 | 1.269 | 0.896 |
| h | 0.71 | 0.096 | 0.069 | 0.052 | 0.083 | 0.077 | 0.057 |
| ξ_w | 0.73 | 0.082 | 0.090 | 0.076 | 0.092 | 0.095 | 0.081 |
| σ_ℓ | 1.92 | 1.652 | 0.640 | 0.532 | 1.506 | 0.939 | 0.831 |
| ξ_p | 0.65 | 0.130 | 0.074 | 0.068 | 0.090 | 0.080 | 0.070 |
| ι_w | 0.59 | 0.205 | 0.165 | 0.159 | 0.190 | 0.168 | 0.160 |
| ι_p | 0.22 | 0.142 | 0.109 | 0.101 | 0.128 | 0.112 | 0.100 |
| ψ | 0.54 | 0.182 | 0.128 | 0.109 | 0.150 | 0.134 | 0.118 |
| r_π | 2.03 | 0.295 | 0.277 | 0.241 | 0.347 | 0.380 | 0.333 |
| ρ | 0.81 | 0.031 | 0.025 | 0.022 | 0.034 | 0.038 | 0.030 |
| r_y | 0.08 | 0.051 | 0.025 | 0.021 | 0.055 | 0.034 | 0.029 |
| $r_{\Delta y}$ | 0.22 | 0.058 | 0.014 | 0.012 | 0.057 | 0.039 | 0.033 |
| ρ_a | 0.95 | 0.071 | 0.028 | 0.020 | - | - | - |
| ρ_b | 0.18 | 0.161 | 0.078 | 0.073 | 0.133 | 0.079 | 0.071 |
| ρ_g | 0.97 | 0.020 | 0.016 | 0.013 | 0.018 | 0.016 | 0.014 |
| ρ_i | 0.71 | - | - | - | - | - | - |
| ρ_r | 0.12 | - | - | - | 0.089 | 0.072 | 0.067 |
| ρ_p | 0.90 | 0.181 | 0.090 | 0.067 | 0.188 | 0.070 | 0.053 |
| ρ_w | 0.97 | 0.031 | 0.030 | 0.019 | 0.022 | 0.029 | 0.021 |
| μ_p | 0.74 | 0.246 | 0.188 | 0.161 | 0.250 | 0.173 | 0.139 |
| μ_w | 0.88 | 0.071 | 0.072 | 0.056 | 0.069 | 0.071 | 0.057 |
| σ_a | 0.45 | 0.441 | 0.061 | 0.052 | - | - | - |
| σ_b | 0.24 | 0.050 | 0.021 | 0.021 | 0.040 | 0.023 | 0.021 |
| σ_g | 0.52 | 0.035 | 0.027 | 0.026 | 0.026 | 0.027 | 0.025 |
| σ_i | 0.45 | - | - | - | - | - | - |
| σ_r | 0.24 | - | - | - | 0.013 | 0.015 | 0.014 |
| σ_p | 0.14 | 0.022 | 0.017 | 0.015 | 0.019 | 0.017 | 0.015 |
| σ_w | 0.24 | 0.026 | 0.021 | 0.020 | 0.022 | 0.023 | 0.021 |

Notes. This table reports the average absolute error across Monte Carlo replications for the indicated parameter and empirical specification. See Table 1 for the definitions of the parameters. The first (second) Monte Carlo experiment corresponds to the case when the true dgp does not include the monetary policy (TFP) disturbance, but the empirical model leaves out the investment disturbance instead.

Table 3: Misspecification tests for the original Smets-Wouters empirical model

| structural SW disturbance excluded | ASD added | marginal data density |
|---|-----------|-----------------------|
| None (original SW) | no | -922.40 |
| TFP, $\varepsilon_{a,t}$ | yes | -931.21 |
| Risk premium, $\varepsilon_{b,t}$ | yes | -908.79 |
| Government expenditure, $\varepsilon_{g,t}$ | yes | -934.14 |
| Investment-specific, $\varepsilon_{i,t}$ | yes | -919.81 |
| Monetary policy, $\varepsilon_{r,t}$ | yes | -926.88 |
| Price mark-up, $\varepsilon_{p,t}$ | yes | -938.85 |
| Wage mark-up, $\varepsilon_{w,t}$ | yes | -947.31 |

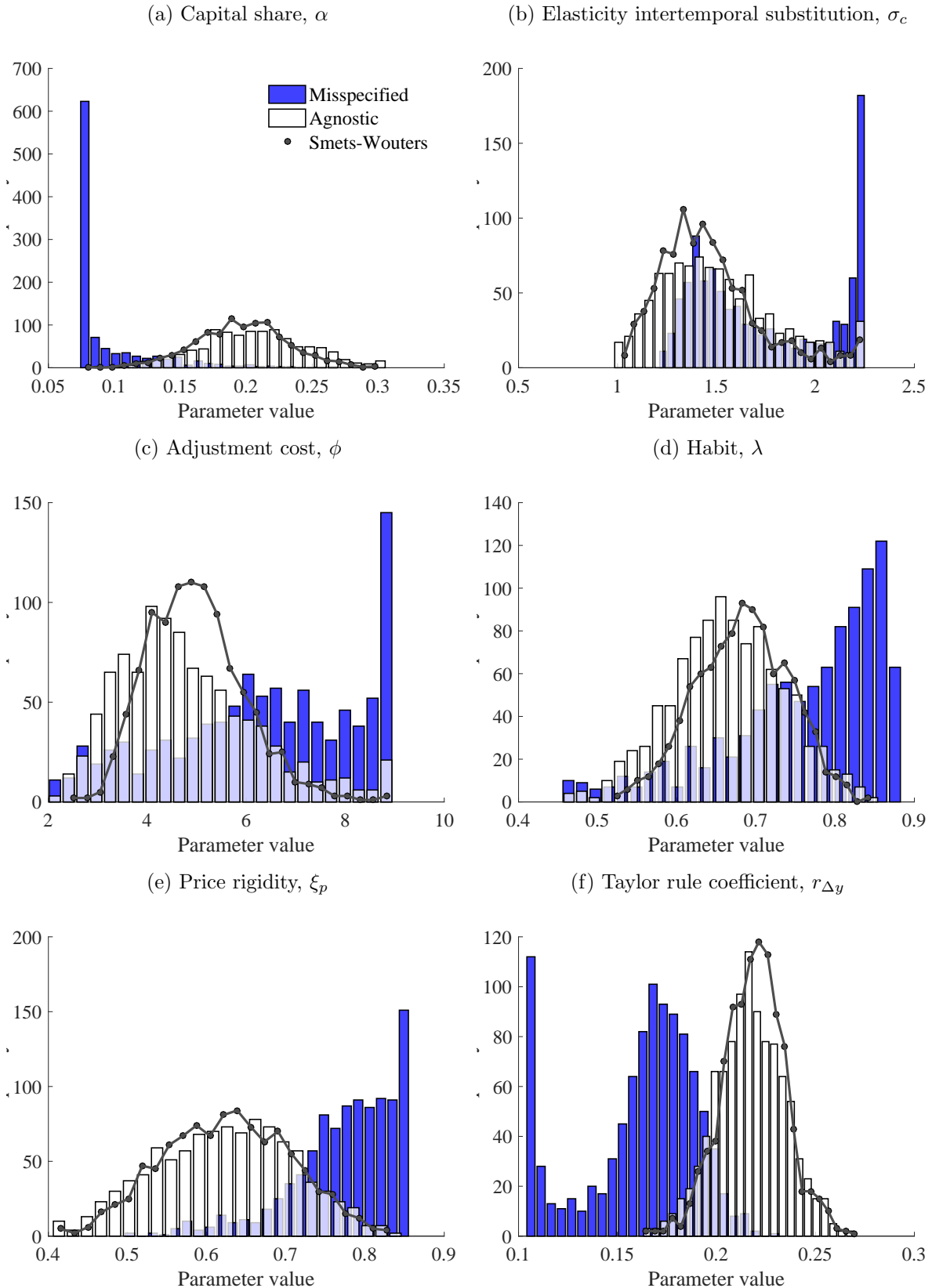
Notes. The table reports the marginal data density for different empirical specifications. The first row reports the value for the original SW specification. The specifications considered in subsequent rows replace the indicated structural disturbance with an agnostic structural disturbance. The bold numbers indicate the cases for which the MDD is higher when the indicated structural disturbance is replaced by an ASD.

Table 4: Role of structural disturbances for variance

| | risk/preference | | investment | | wage |
|--------------------|------------------------|--------------------------------------|------------------------|--------------------------------------|--------------------------------------|
| | SW $\varepsilon_{b,t}$ | agnostic $\tilde{\varepsilon}_{A,t}$ | SW $\varepsilon_{i,t}$ | agnostic $\tilde{\varepsilon}_{B,t}$ | agnostic $\tilde{\varepsilon}_{C,t}$ |
| output | 1.53 | 1.14 | 7.34 | 2.17 | 0.28 |
| flex. price output | 0 | 2.08 | 5.39 | 1.02 | 0.36 |
| consumption | 2.18 | 1.51 | 2.83 | 0.49 | 0.25 |
| investment | 0.22 | 1.06 | 44.2 | 29.3 | 1.00 |
| hours | 2.52 | 1.29 | 8.15 | 4.97 | 2.03 |
| capital | 0.04 | 0.12 | 32.5 | 2.37 | 9.75 |
| utilization | 0.86 | 4.14 | 35.4 | 9.46 | 14.7 |
| price of capital | 45.4 | 18.6 | 36.0 | 31.6 | 7.21 |
| marginal cost | 0.87 | 15.2 | 3.11 | 2.61 | 5.13 |
| policy rate | 7.40 | 17.2 | 18.3 | 12.5 | 0.65 |
| inflation | 0.58 | 0.68 | 3.18 | 3.96 | 0.91 |
| output growth | 22.1 | 21.3 | 15.8 | 8.04 | 1.82 |
| consumption growth | 61.2 | 61.7 | 0.95 | 2.03 | 0.10 |
| investment growth | 2.46 | 12.6 | 82.1 | 70.0 | 0.81 |

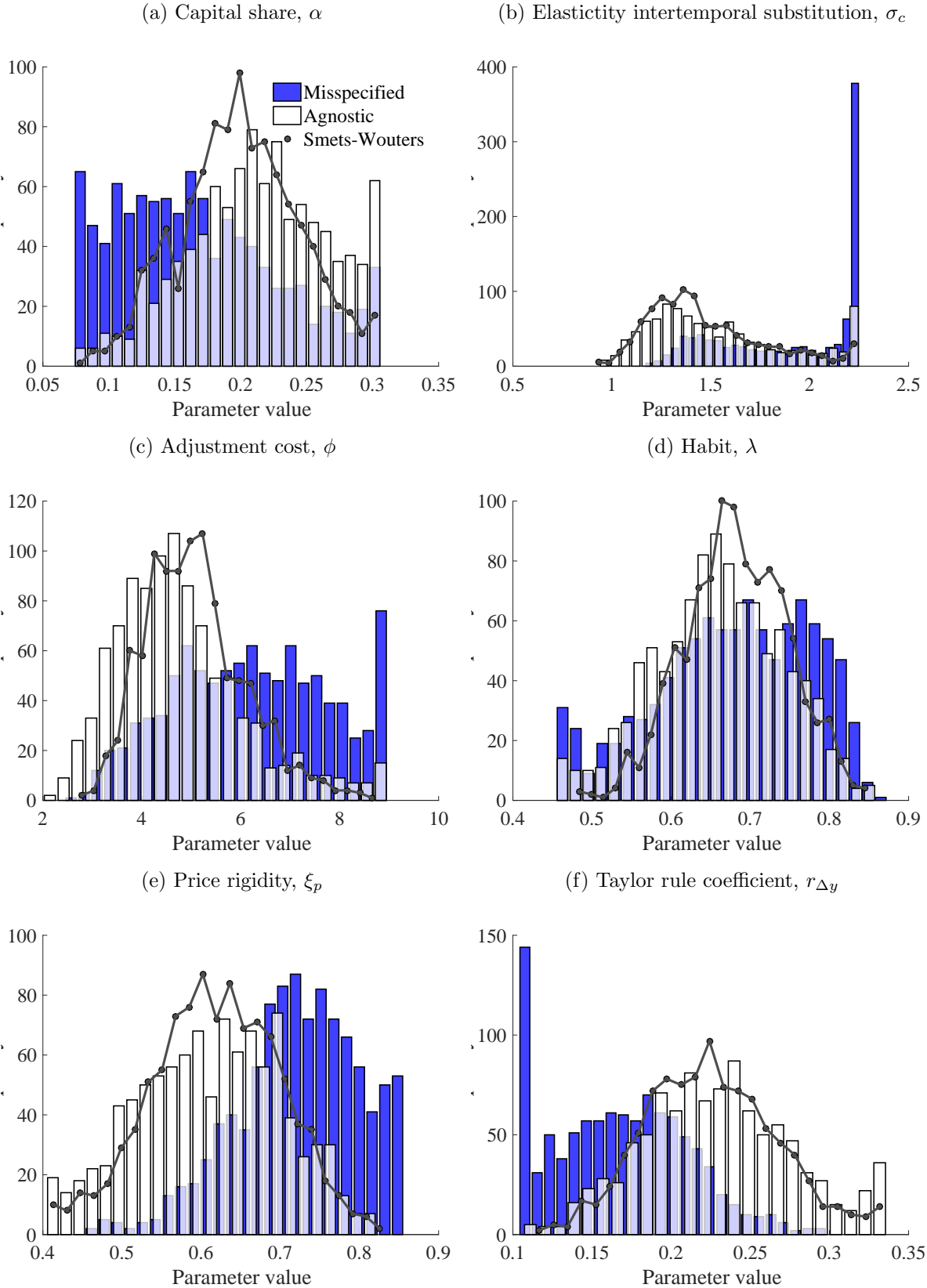
Notes. The table reports the percentage of total variability explained by the SW versus the agnostic risk-premium disturbance, the SW versus the agnostic investment disturbance, and the agnostic wage disturbance. The numbers for the SW disturbance are from estimation of the original SW model. The numbers for the agnostic disturbance are from our preferred empirical model with three ASDs.

Figure 1: Histograms for parameter estimates: First Monte Carlo experiment



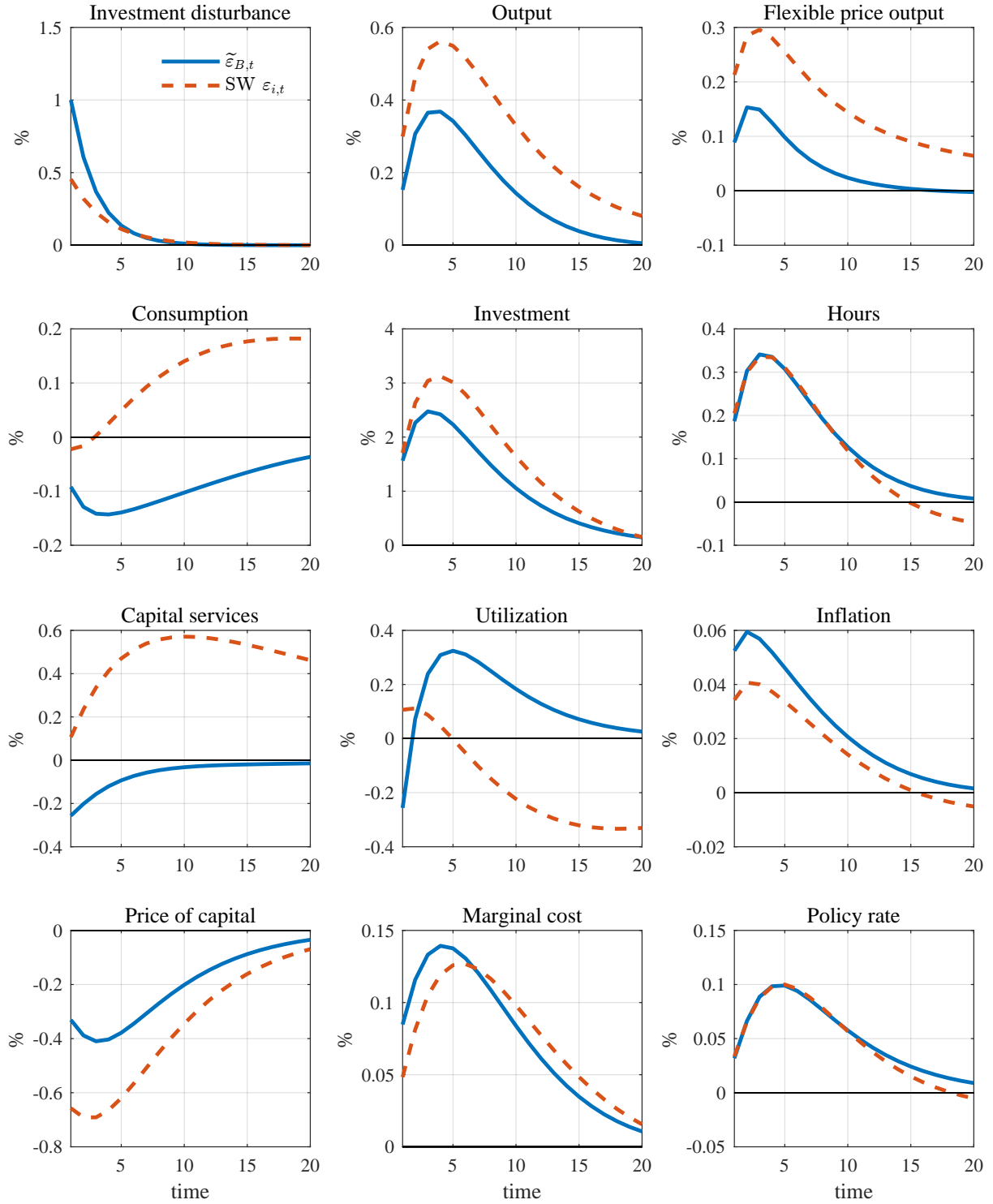
Notes. The panels plot the distribution of the indicated parameter across the Monte Carlo replications. The color of the histograms for the misspecified case changes in a lighter shade when they overlap with the histogram for the agnostic specification. In this experiment, the true *dgp* does not include the monetary policy disturbance, but the empirical model leaves out the investment disturbance instead.

Figure 2: Histograms for parameter estimates: Second Monte Carlo experiment



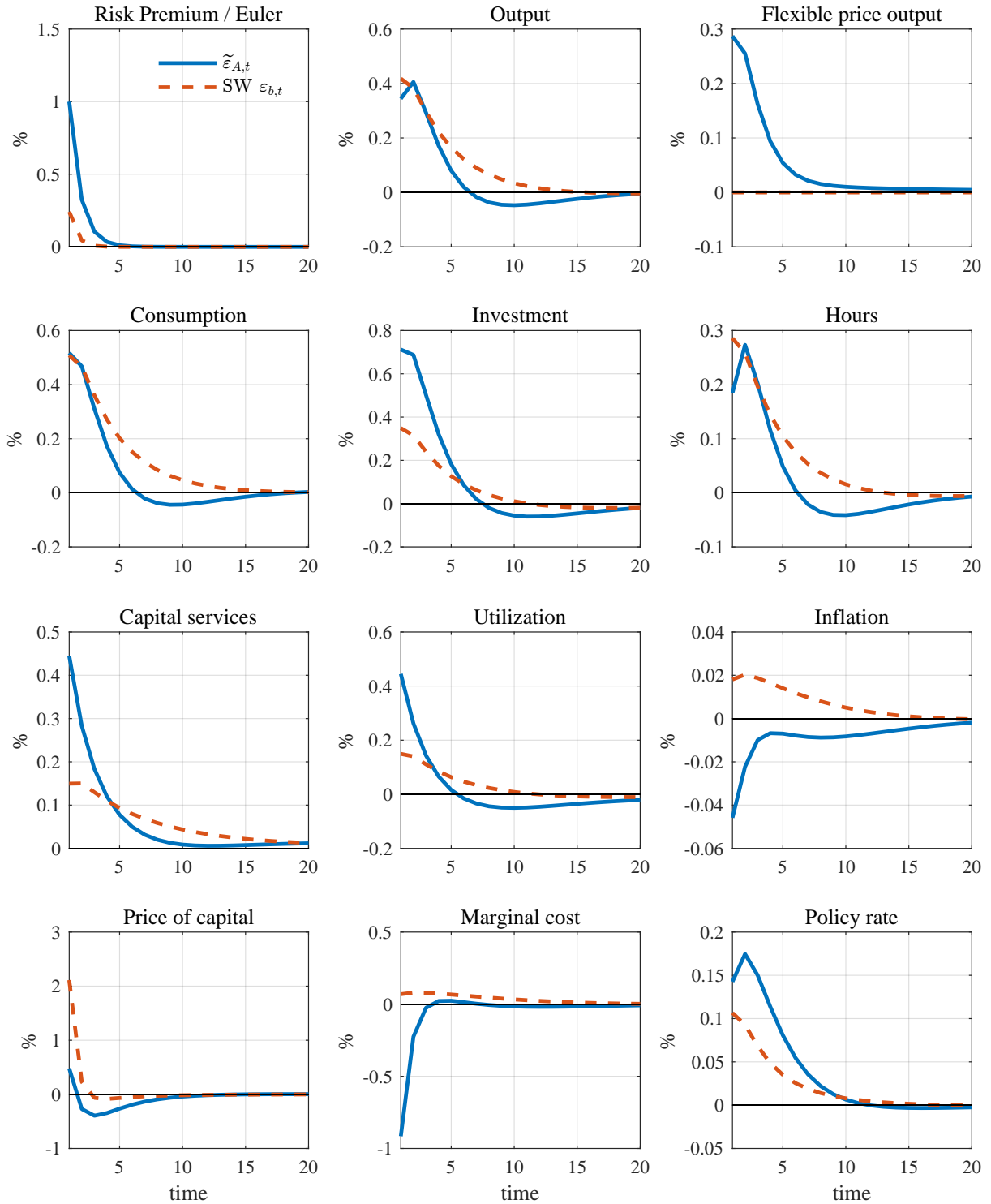
Notes. The panels plot the distribution of the indicated parameter across the Monte Carlo replications. The color of the histograms for the misspecified case changes in a lighter shade when they overlap with the histogram for the agnostic specification. In this experiment, the true *dgp* does not include the TFP disturbance, but the empirical model leaves out the investment disturbance instead.

Figure 3: IRFs of the SW investment and the agnostic investment-modernization disturbance



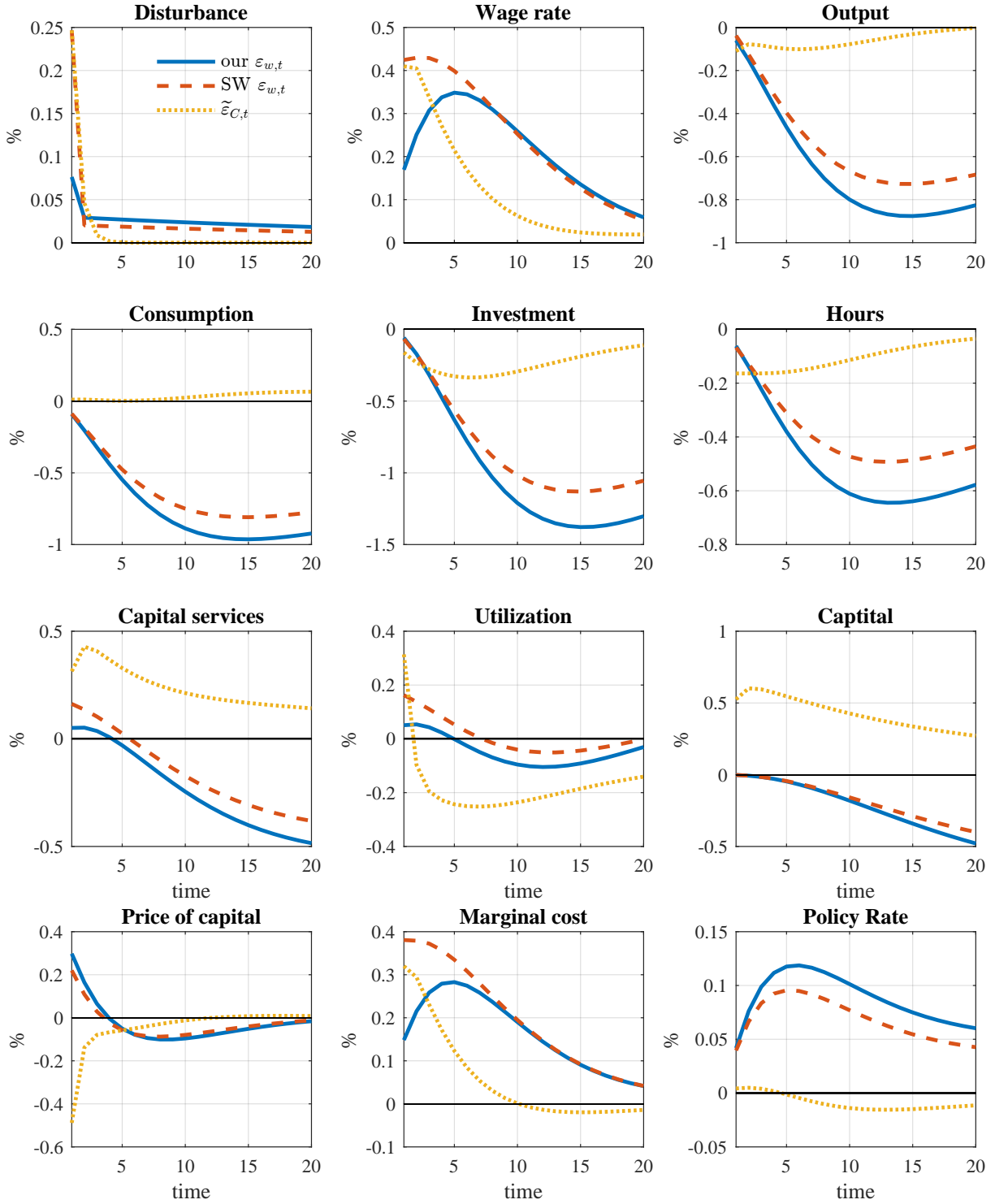
Notes. These panels plot the IRFs of the SW investment-specific productivity disturbance and the agnostic disturbance $\tilde{\varepsilon}_{B,t}$ that we interpret as an investment-modernization disturbance.

Figure 4: IRFs of the SW risk-premium and the agnostic Euler disturbance



Notes. These panels plot the IRFs of the SW risk-premium disturbance and the agnostic disturbance $\tilde{\varepsilon}_{A,t}$ that we interpret as an Euler disturbance.

Figure 5: IRFs of the agnostic capital-efficiency wage mark-up disturbance



Notes. These panels plot the IRFs of the agnostic disturbance $\tilde{\varepsilon}_{C,t}$ that we interpret as a capital-efficiency wage mark-up disturbance. They also plot the SW wage disturbance for the original SW specification and ours.

610 **References**

- 611 An, S., Schorfheide, F., 2007. Bayesian analysis of dsge models. *Econometric Reviews* 26,
612 113–172.
- 613 Ascari, G., Phaneuf, L., Sims, E., 2016. Business cycles, investment shocks, and the "barro-
614 king curse". Department of Economics Discussion Paper Series, Number 815.
- 615 Beltran, D.O., Draper, D., 2015. Estimating dynamic macroeconomic models: How informa-
616 tive are the data? Unpublished manuscript.
- 617 Campbell, J.R., 1998. Entry, exit, embodied technology, and business cycles. *Review of*
618 *Economic Dynamics* 1, 371–408.
- 619 Canova, F., Ferroni, F., Matthes, C., 2015. Approximating time varying structural models
620 with time invariant structures. Unpublished manuscript.
- 621 Chari, V.V., Kehoe, P.J., McGrattan, E.R., 2007. Business cycle accounting. *Econometrica*
622 Forthcoming.
- 623 Chari, V.V., Kehoe, P.J., McGrattan, E.R., 2008. Are structural vars with long-run restric-
624 tions useful in developing business cycle theory. *Journal of Monetary Economics* 55.
- 625 Cúrdia, V., Reis, R., 2012. Correlated disturbances and u.s. business cycles. NBER Working
626 Paper No. 15774.
- 627 Del Negro, M., Schorfheide, F., 2009. Monetary policy analysis with potentially misspecified
628 models. *American Economic Review* 99, 1415–1450.
- 629 Del Negro, M., Schorfheide, F., Smets, F., Wouters, R., 2007. On the fit of new keynesian
630 models. *Journal of Business and Economic Statistics* 25, 123–143.
- 631 Den Haan, W.J., Levin, A., 1997. A practitioner's guide to robust covariance matrix estimation,
632 in: *Handbook of Statistics* 15. Elsevier, Holland, pp. 291–341.
- 633 Ferroni, F., Grassi, S., León-Ledesma, M.A., 2015. Fundamental shock selection in dsge
634 models. Unpublished manuscript.
- 635 Galizia, D., 2015. Misspecification and the causes of business cycles. Unpublished manuscript.
- 636 Gertler, M., Karadi, P., 2010. A model of unconventional monetary policy. *Journal of*
637 *Monetary Economics* 58, 17–34.
- 638 Inoue, A., Kuo, C.H., Rossi, B., 2015. Identifying the sources of model misspecification.
639 Unpublished manuscript, Pompeu Fabra,.
- 640 Ireland, P., 2004. A method for taking models to the data. *Journal of Economics Dynamics*
641 *and Control* 28, 1205–1226.

- 642 Iskrev, N., 2010. Local identification in dsge models. *Journal of Monetary Economics* 57,
643 189–202.
- 644 Justiano, A., Primiceri, G., Tambalotti, A., 2010. Investment shocks and business cycles.
645 *Journal of Monetary Economics* 57, 132–145.
- 646 Komunjer, I., Ng, S., 2011. Dynamic identification of dynamic stochastic general equilibrium
647 models. *Econometrica* 79, 1995–2032.
- 648 Mickelsson, G., 2015. Estimation of dsge models: Maximum likelihood vs. bayesian methods.
649 Uppsala University Department of Economics Working Paper 2015:6.
- 650 Onatski, A., Williams, N., 2010. Empirical and policy performance of a forward-looking
651 monetary model. *Journal of Applied Econometrics* 25, 145–176.
- 652 Paccagnini, A., 2017. Dealing with misspecification in dsge models: A survey. Munich
653 Personal RePec Archive Paper No. 82914.
- 654 Smets, F., Wouters, R., 2003. An estimated stochastic dynamic general equilibrium model
655 of the euro area. *Journal of the European Economic Association* 1, 1123–1175.
- 656 Smets, F., Wouters, R., 2007. Shocks and frictions in us business cycles: A bayesian dsge
657 approach. *American Economic Review* 97, 586–606.
- 658 Stock, J.H., Watson, M.W., 1999. Business cycle fluctuations in us macroeconomic time
659 series, in: Taylor, J.B., Woodford, M. (Eds.), *Handbook of Macroeconomics*. Elsevier,
660 Amsterdam, The Netherlands. *Handbook of Macroeconomics*, pp. 3–64.

661 Appendix A. The ASD procedure: General setup

662 Using a simple model, Section 2 made clear that adding an agnostic structural disturbance
 663 (ASD) is as simple as adding an exogenous random variable multiplied by a coefficient to *each*
 664 model equation. The coefficients are unrestricted and not related to any of the structural
 665 parameters of the model. It also shows how the ASD coefficients appear in the model’s policy
 666 functions and that an alternative way to incorporate ASDs consists of adding them directly
 667 to the policy functions. In this appendix, we generalize the discussion.

668 Appendix A.1. Adding ASDs to model equations

Consider the following linearized model:

$$0_{N \times 1} = \mathbb{E}_t [\Lambda_2(\Psi) s_{t+1} + \Lambda_1(\Psi) s_t + \Lambda_0(\Psi) s_{t-1} + \Gamma(\Psi) \varepsilon_{t+1} + \Upsilon(\Psi) \varepsilon_t], \quad (\text{A.1a})$$

$$\varepsilon_t = P\varepsilon_{t-1} + \Omega\eta_t, \quad (\text{A.1b})$$

$$\mathbb{E}_t [\eta_{t+1}] = 0, \quad (\text{A.1c})$$

$$\mathbb{E}_t [\eta_{t+1}\eta'_{t+1}] = I_{M \times M}, \quad (\text{A.1d})$$

669 where Ψ is the vector containing the structural parameters, s_t is the $N \times 1$ vector of en-
 670 dogenous variables, and ε_t is the $M \times 1$ vector of exogenous random variables. All variables
 671 are defined relative to their steady state values. Most linearized DSGE models can be rep-
 672 resented with such a system of equations. The literature typically assumes that innovations
 673 are assumed to be orthogonal to each other, that is, Ω is assumed to be a diagonal matrix.
 674 We make the same assumption.²⁷

We first discuss the case for which s_t includes only state variables and all N state variables are observables. Suppose that the researcher is only sure about M_1 structural disturbances. These are part of the vector, $\varepsilon_{1,t}$. If $M_1 < N$ and there are no other disturbances, then there is a singularity problem. One option would be to add measurement error. But structural disturbances and measurement errors are very different. Structural disturbances affect economic variables and propagate through the system according to the economic mechanisms of the model. As discussed in the main text, measurement error disturbances do not. Moreover, most researchers would find it undesirable if measurement error “explains” a large part of the data. Another option is to make a best guess and to add a vector $\varepsilon_{2,t}$ with M_2 additional

²⁷The literature typically also assumes that P is diagonal. An exception is Cúrdia and Reis (2012).

structural disturbances with $M_2 \geq N - M_1$. Equation (A.1) can then be written as

$$0_{N \times 1} = \mathbb{E}_t [\Lambda_2(\Psi) s_{t+1} + \Lambda_1(\Psi) s_t + \Lambda_0(\Psi) s_{t-1} + \Gamma(\Psi) \varepsilon_{t+1} + \Upsilon(\Psi) \varepsilon_t]$$

$$= \mathbb{E}_t \left[\begin{array}{c} \Lambda_2(\Psi) s_{t+1} + \Lambda_1(\Psi) s_t + \Lambda_0(\Psi) s_{t-1} \\ + [\Gamma_1(\Psi) \quad \Gamma_2(\Psi)] \begin{bmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \end{bmatrix} + [\Upsilon_1(\Psi) \quad \Upsilon_2(\Psi)] \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \end{array} \right], \quad (\text{A.2a})$$

$$\begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix} + \begin{bmatrix} \Omega_{11} & 0_{M_1 \times M_2} \\ 0_{M_2 \times M_1} & \Omega_{22} \end{bmatrix} \begin{bmatrix} \eta_{1,t} \\ \eta_{2,t} \end{bmatrix}, \quad (\text{A.2b})$$

$$\mathbb{E}_t \begin{bmatrix} \eta_{1,t+1} \\ \eta_{2,t+1} \end{bmatrix} = 0, \quad (\text{A.2c})$$

$$\mathbb{E}_t \left[\begin{bmatrix} \eta_{1,t+1} \\ \eta_{2,t+1} \end{bmatrix} \begin{bmatrix} \eta_{1,t+1} & \eta_{2,t+1} \end{bmatrix} \right] = I_{M \times M}. \quad (\text{A.2d})$$

675 The column vectors $\Gamma_2(\Psi)$ and $\Upsilon_2(\Psi)$ capture the restrictions imposed by the M_2 additional
676 structural disturbances.

Introducing ASDs. With ASDs the system of equations is modified as follows:

$$0_{N \times 1} = \mathbb{E}_t \left[\begin{array}{c} \Lambda_2(\Psi) s_{t+1} + \Lambda_1(\Psi) s_t + \Lambda_0(\Psi) s_{t-1} \\ + \Gamma_1(\Psi) \varepsilon_{1,t+1} + \Upsilon_1(\Psi) \varepsilon_{1,t} + \tilde{\Upsilon}_2 \tilde{\varepsilon}_{2,t} \end{array} \right]. \quad (\text{A.3})$$

677 As in the main text, all ASD variables and their associated coefficients are denoted with a
678 tilde. As long as $P_{21} = 0$, then one can exclude $\tilde{\varepsilon}_{2,t+1}$ from the system, since what matters is
679 $\mathbb{E}_t[\tilde{\varepsilon}_{2,t+1}] = P_{22}\tilde{\varepsilon}_{2,t}$ and the reduced-form coefficient $\tilde{\Upsilon}_2$ captures this forward looking aspect
680 of ASDs as well as the contemporaneous impact of the ASD on the model equations.²⁸
681 Thus, adding an agnostic disturbance introduces one additional parameter for each model
682 equation.^{29,30} Replacing regular structural disturbances with agnostic structural disturbances
683 may make it harder to identify Ψ , the structural parameters of the model. As discussed in
684 Appendix C.2, this turned out to be not an issue for the experiments discussed in this paper.

685 Appendix A.2. Adding ASDs to model solutions

686 We start this section with a proposition that will be helpful with the second formulation
687 of the ASD procedure. Consider again the model given in Equation (A.2), which divides the
688 vector with exogenous disturbances, ε_t , into two parts, the $M_1 \times 1$ vector, $\varepsilon_{1,t}$, and the $M_2 \times 1$

²⁸If $P_{2,1} \neq 0$, then additional reduced-form coefficients would be needed to capture the predictive power of $\varepsilon_{1,t}$ for future values of $\tilde{\varepsilon}_{2,t}$. The assumption that $P_{2,1} = 0$ is weak given that the standard assumption in the literature is that structural disturbances are independent random variables.

²⁹Without loss of generality one can set the standard deviations of the innovation of the ASDs equal to 1, which in this case is a normalization of the diagonal elements of $\Omega_{2,2}$. As with regular structural disturbances, one would need to estimate the parameters of the time series specification contained in G .

³⁰As discussed in Appendix E, one could choose to leave the agnostic disturbance out of some equations.

689 vector, $\varepsilon_{2,t}$. A recursive solution to Equation (A.2) has the following form:

$$\begin{aligned} s_t &= A(\Psi) s_{t-1} + B(\Psi) \varepsilon_t \\ &= A(\Psi) s_{t-1} + \begin{bmatrix} B_1(\Psi) & B_2(\Psi) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}. \end{aligned} \quad (\text{A.4})$$

690 The following proposition states that the properties of $\varepsilon_{2,t}$ do *not* affect the coefficients
691 of the policy rule related to s_{t-1} and $\varepsilon_{1,t}$, that is, they do not affect $A(\Psi)$ and $B_1(\Psi)$. Thus,
692 it does not matter whether $\varepsilon_{2,t}$ is a regular or an agnostic structural disturbances and the
693 time series properties of $\varepsilon_{2,t}$ do not matter either. The only assumption needed is that the
694 elements of P_{21} are equal to zero, which corresponds to the case when $\varepsilon_{1,t}$ has no effect on
695 *future* values of $\varepsilon_{2,t}$. This is not very restrictive given that the literature usually sets all
696 elements of P_{21} equal to zero (and also all elements of P_{12} , $\Omega_{1,2}$, and $\Omega_{2,2}$ as well as the
697 off-diagonal elements of P_{11} , P_{22} , $\Omega_{1,1}$ and $\Omega_{2,2}$).

698 **Proposition 1.** *If the model is given by equation (A.2) and all elements of P_{21} are equal to*
699 *zero, then (i) $A(\Psi)$ and $B_1(\Psi)$ do not depend on $\Gamma_2(\Psi)$ and $\Upsilon_2(\Psi)$, which characterize the*
700 *nature of the additional disturbances, and (ii) $A(\Psi)$ and $B_1(\Psi)$ do not depend on P_{22} , Ω_{21} ,*
701 *and Ω_{22} , which characterize the time series properties of $\varepsilon_{2,t}$.*

702 **Proof.** Substitution of the policy rule as given in Equation (A.4) into the system of Equa-
703 tions (A.2) gives,

$$0_{N \times 1} = (\Lambda_2 A^2 + \Lambda_1 A + \Lambda_0) s_{t-1} + (\Lambda_2 AB + \Lambda_2 BP + \Lambda_1 B + \Gamma P + \Upsilon) \varepsilon_t, \quad (\text{A.5})$$

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{1,t} & \varepsilon_{2,t} \end{bmatrix}', \quad (\text{A.6})$$

$$B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}, \quad (\text{A.7})$$

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}, \quad (\text{A.8})$$

where we have suppressed the dependence of coefficients on Ψ . The first equation has to hold for all values of s_{t-1} and ε_t . This implies that a solution must satisfy

$$\Lambda_2 A^2 + \Lambda_1 A + \Lambda_0 = 0_{N \times N} \quad (\text{A.9})$$

and

$$\Lambda_2 AB + \Lambda_2 BG + \Lambda_1 B + \Gamma P + \Upsilon = 0_{N \times (M_1 + M_2)}. \quad (\text{A.10})$$

A does not depend on the time series properties of $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, since B , P , and Ω do not appear in Equation (A.9). Equation (A.10) can be written as follows

$$\bar{\Lambda} \begin{bmatrix} B_1 & B_2 \end{bmatrix} + \Lambda_2 \begin{bmatrix} B_1 & B_2 \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} + \Gamma \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} + \Upsilon = 0_{N \times (M_1 + M_2)}, \quad (\text{A.11})$$

704 where $\bar{\Lambda} = \Lambda_2 A + \Lambda_1$. This is a system of $N \times (M_1 + M_2)$ equations to solve for the elements
 705 of B . It can be split into the following two sets of systems:

$$\bar{\Lambda} B_1 + \Lambda_2 B_1 P_{11} + \Lambda_2 B_2 P_{21} + \Gamma_1 \begin{bmatrix} P_{11} \\ P_{21} \end{bmatrix} + \Upsilon_1 = 0_{N \times M_1}, \quad (\text{A.12})$$

$$\bar{\Lambda} B_2 + \Lambda_2 B_1 P_{12} + \Lambda_2 B_2 P_{22} + \Gamma_2 \begin{bmatrix} P_{12} \\ P_{22} \end{bmatrix} + \Upsilon_2 = 0_{N \times M_2}. \quad (\text{A.13})$$

706 If $P_{21} = 0$, then Equation (A.12) contains $N \times M_1$ equations to solve for all the elements
 707 of B_1 . The solution cannot depend on P_{22} or Ω_{22} since these matrices do not appear in this
 708 equation. ■

709 It is intuitive that the elements of P_{21} have to be equal to zero, that is, $\varepsilon_{1,t}$ should not
 710 affect future values of $\varepsilon_{2,t}$. If current values of $\varepsilon_{1,t}$ do affect future values of $\varepsilon_{2,t}$ and therefore
 711 future values of s_t , then one has to know how $\varepsilon_{2,t}$ affects model outcomes to determine how
 712 $\varepsilon_{1,t}$ affects current outcomes for s_t .

713 **Introducing ASDs.** The solution to a linearized model can be written as:

$$s_t = \sum_{m=1}^M s_t^{[m]}, \quad (\text{A.14})$$

$$s_t^{[m]} = A(\Psi) s_{t-1}^{[m]} + B_{.,m}(\Psi) \varepsilon_{m,t}, \quad (\text{A.15})$$

714 where $s_t^{[m]}$ represents the outcome of the state variable if the *only* disturbance in the economy
 715 is the m^{th} -disturbance, $\varepsilon_{m,t}$, and $B_{.,m}$ is the m^{th} column of B . Thus, one can think of the s_t
 716 variables as the sum of the outcomes in “one-disturbance” economies. The linearity of the
 717 model is important for this additive property. According to Proposition 1, the coefficients
 718 on the lagged state variable, $A(\Psi)$, do not depend on the particular disturbance considered.
 719 That is, whereas $B_{.,m}(\Psi)$ is indexed by m because it depends on what kind of disturbance
 720 is the driving force of the economy, $A(\Psi)$ does not. This property greatly increases the
 721 efficiency of our procedure.

722 Our proposed procedure consists of including \widetilde{M}_2 agnostic structural disturbances. This
 723 results in the following time series representation of the policy functions:³¹

$$s_t = \sum_{m=1}^M s_t^{[m]}, \quad (\text{A.16a})$$

$$s_t^{[m]} = A(\Psi) s_{t-1}^{[m]} + B_{.,m}(\Psi) \varepsilon_{m,t} \quad \text{for } m \leq M_1, \quad (\text{A.16b})$$

$$s_t^{[m]} = A(\Psi) s_{t-1}^{[m]} + \widetilde{B}_{.,m} \widetilde{\varepsilon}_{m,t} \quad \text{for } M_1 + 1 \leq m \leq M_1 + \widetilde{M}_2 = M. \quad (\text{A.16c})$$

³¹Proposition 1 indicates that this specification is valid as long as the elements of P_{12} are equal to zero, which is usually the case.

724 In terms of notation, $B_{.,m}(\Psi)$ contains coefficients associated with a regular structural dis-
725 turbance which are a function of Ψ and $\tilde{B}_{.,m}$ contains reduced-form coefficients associated
726 with a structural agnostic disturbance. The *only* difference between this specification and the
727 standard DSGE specification with only regular structural disturbances is that the $\tilde{B}_{.,m}$ coef-
728 ficients are unrestricted reduced-form coefficients. Since our agnostic disturbances are struc-
729 tural disturbances, their impact propagates through the system exactly as regular structural
730 disturbances do, that is, as described by $A(\Psi)$. The property of linear models that $A(\Psi)$ does
731 not depend at all on what is the nature of the structural disturbances nor on their time series
732 properties makes it possible to efficiently add structural disturbances to the specification
733 without having to be specific on what they are.

734 The dimension of $\tilde{B}_{.,m}$ is equal to N , the number of state variables. This means that
735 adding an agnostic disturbance means estimating an additional N parameters. The number
736 of additional parameters to be estimated is limited because structural disturbances differ
737 in their initial impact, but their propagation through time is the same for all disturbances
738 and controlled by $A(\Psi)$. Moreover, an increase in the standard deviation of an agnostic
739 structural disturbance affects the model variables in exactly the same way as an identical
740 proportional increase of the elements of $\tilde{B}_{.,m}$. Consequently, the standard deviation of an
741 agnostic disturbance can be normalized to equal 1.³²

742 **Adding observation equations.** If there are observables that are not state variables,
743 then one also needs additional equations for these y_t variables, which for our set-up is given
744 by

$$y_t = \sum_{m=1}^M y_t^{[m]}, \quad (\text{A.17a})$$

$$y_t^{[m]} = C(\Psi) s_{t-1}^{[m]} + D_{.,m}(\Psi) \varepsilon_{m,t} \quad \text{for } m \leq M_1, \quad (\text{A.17b})$$

$$y_t^{[m]} = C(\Psi) s_{t-1}^{[m]} + \tilde{D}_{.,m} \tilde{\varepsilon}_{m,t} \quad \text{for } M_1 + 1 \leq m \leq M_1 + \tilde{M}_2 = M, \quad (\text{A.17c})$$

745 where y_t is the $(\bar{N} \times 1)$ vector with observables that are not state variables. Each additional
746 observable used in the estimation will introduce one more coefficient related to the agnostic
747 structural disturbances.

748 **Unobserved components.** The system represented in Equation (A.16) makes clear that
749 ASDs can be interpreted simply as unobserved components that are added to a theoretical
750 block. Using reduced-form systems with unobserved components is common practice in
751 macroeconomic time-series models. The data will provide information about the $\tilde{B}_{.,m}$ (and
752 $\tilde{D}_{.,m}$) coefficients. If one mainly cares about getting good time-series representations of

³²If the time series processes of the two disturbances have the same number of parameters, then replacing a regular structural disturbance by an agnostic disturbance typically means estimating an additional $N - 1$ parameters.

753 macroeconomic variables, then can just use this way of incorporating ASDs. However, as
754 shown in our empirical applications, ASDs become much more interesting if one knows the
755 $\tilde{\Upsilon}_{2,m}$ coefficients. Then one can much better understand how these unobserved components
756 affect the model economy.

757 **Appendix B. Misspecification: Literature review**

758 Most empirical papers that estimate a dynamic macroeconomic model do not raise the
759 issue of model uncertainty or misspecification, except possibly with some robustness exer-
760 cises.³³ This does – of course – not mean that the profession is not aware that misspecification
761 is a serious concern. In fact, some of the most prominent researchers in this research area have
762 drawn attention to the risk of misspecification. The first subsection discusses evidence that
763 indicates that misspecification of DSGE models is a serious concern. The second subsection
764 discusses approaches proposed in the literature to deal with misspecification. See Paccagnini
765 (2017) for a more detailed survey.

766 **Appendix B.1. Indications of DSGE misspecification**

767 Del Negro et al. (2007) develop a procedure that allows the data to determine the use-
768 fulness of a DSGE model relative to a much less restricted VAR. Using a model very similar
769 to the DSGE model of Smets and Wouters (2003), they find that their procedure does put
770 some weight on the DSGE model, which implies that the restrictions of the DSGE model are
771 of some value. However, they also argue that misspecification is a concern that “... *is not*
772 *small enough to be ignored.*” Using the same methodology, Del Negro and Schorfheide (2009)
773 also find “... *strong evidence of DSGE model misspecification.*”

774 There is also more indirect evidence that misspecification of estimated DSGE models is
775 substantive. Using the Smets and Wouters (2003) model for the Euro Area, Beltran and
776 Draper (2015) find that the data prefer implausible estimates for several parameters. For
777 example, *most* of the mass of the marginal likelihood for the parameter of relative risk aversion
778 is above 200, way above the range of values considered reasonable. This information provided
779 by the likelihood is typically not revealed in empirical studies, since only properties of the
780 posterior are reported and the choice of prior ensures that these aspects of the empirical
781 likelihood have little or no weight in the posterior. A similar conclusion can be drawn from
782 Onatski and Williams (2010). They estimate the same model using uniform priors over
783 bounded ranges. These ranges are such that the priors are less informative than the ones
784 typically used in the literature. Consistent with the results in Beltran and Draper (2015),
785 several of the point estimates in Onatski and Williams (2010) are at the prior bounds. Using
786 a new algorithm to deal with the complexity of estimating likelihood functions, Mickelsson
787 (2015) re-estimates the model of Smets and Wouters (2007) and he also finds that several

³³Interestingly, there are quite a few macroeconomic models in which agents – especially agents setting fiscal and monetary policy – face model uncertainty. If policy makers face model uncertainty about the correct model, then researchers are likely to do so as well.

788 parameter estimates are significantly different from the ones reported in Smets and Wouters
789 (2007).

790 Another possible reason for misspecification is the assumption that parameters are con-
791 stant. To get efficient estimates we would like to use long time-series data, but the longer
792 the time series the less likely that all parameters are constant. Canova et al. (2015) address
793 this issue and document that this is important for the model of Gertler and Karadi (2010).³⁴

794 **Appendix B.2. Dealing with misspecification: Other approaches**

795 **Richer models.** Exogenous random disturbances are typically assumed not to be corre-
796 lated with each other. This is a convenient assumption, because allowing for interaction
797 between the different exogenous disturbances would substantially increase the number of pa-
798 rameters to be estimated given that DSGE typically have a several exogenous disturbances.
799 However, it seems quite plausible that such disturbances are correlated. Del Negro and
800 Schorfheide (2009) and Cúrdia and Reis (2012) deal with this possible misspecification and
801 allow for more general processes to describe the behavior of the exogenous random distur-
802 bances.

803 Cúrdia and Reis (2012) find that this generalization has nontrivial consequences for the
804 properties of the model. For example, the impact of a monetary policy shock on output
805 is only half as big when the exogenous random variables are allowed to be correlated and
806 the medium-term impact of a government spending shock switches from being positive to
807 negative.³⁵

808 Enriching a model by allowing for additional features and more general specifications is
809 likely to reduce misspecification. However, richer models typically have more parameters,
810 which will reduce the efficiency of the estimation by reducing the number of degrees of
811 freedom.

812 **Multiple models.** Another way to deal with potential misspecification is to consider a
813 set of different DSGE models. These could be compared informally or formally using, for
814 example, relative marginal likelihoods or model averaging.³⁶ However, given the difficulty
815 of modeling macroeconomic phenomena, it seems likely that *all* models in a set of DSGE
816 models are subject to at least some type of misspecification.

817 **Combining structural and reduced-form models.** Ireland (2004) is an early paper that
818 proposes a more general procedure to deal with possible misspecification when estimating a
819 DSGE model even though the word misspecification is not used in the paper. Specifically,
820 Ireland (2004) “... *augments the DSGE model so that its residuals – meaning the movements*

³⁴The literature cited in Canova et al. (2015) documents that this is an issue in a variety of DSGE models.

³⁵Cúrdia and Reis (2012) still impose the standard assumption that the *innovations* of the shocks are uncorrelated.

³⁶See chapter 5 in An and Schorfheide (2007) for a detailed discussion.

821 *in the data that the theory cannot explain – are described by a VAR.*” To understand this
 822 procedure, consider the following representation of the linearized solution of a DSGE model:

$$s_t = As_{t-1} + B\eta_t, \tag{B.1}$$

$$y_t = Cs_{t-1} + D\eta_t, \tag{B.2}$$

where s_t is a vector containing (endogenous and exogenous) state variables, y_t is a vector containing the observables, and η_t is a vector containing the innovations of the exogenous random variables. Ireland (2004) proposes to augment the observation Equation (B.2) as follows:

$$y_t = Cs_{t-1} + D\eta_t + u_t \tag{B.3a}$$

$$u_t = Fu_{t-1} + \xi_t \tag{B.3b}$$

823 where u_t captures the misspecification or incompleteness of the DSGE model. In his appli-
 824 cation, the structural equations are the policy rules from a standard Real Business Cycle
 825 (RBC) model with total factor productivity (TFP) as the only driving process. If the stan-
 826 dard deviation of η_t is equal to 0, then this procedure boils down to estimating a standard
 827 VAR.

828 Note that the presence of the “missing elements” that are captured by u_t is assumed to
 829 have no effect on that part of agents’ behavior that is described by the DSGE model, that is,
 830 the matrices A , B , C , and D . For this to be correct it must be true that the response of the
 831 economy to a TFP shock does not depend on the presence of other disturbances. One might
 832 think that such independence of a DSGE’s policy rule to the presence of other disturbances
 833 is only correct if the additional disturbances represent measurement error.³⁷ However, this
 834 “independence” property is correct in linear(ized) models in the sense that the specification
 835 of the structural part given in Equations (B.1) and (B.2) does not depend on the presence of
 836 not included structural disturbances. It must be noted that the assumption that u_t follows a
 837 first-order (or even a finite-order) VAR could very well be restrictive. Thus the reduced-form
 838 specification for u_t could be misspecified as well.

The most comprehensive methodology to deal with misspecified DSGE models is put forward in Del Negro et al. (2007). Their starting point is a VAR specification of the observables. That is,

$$y_t = \sum_{k=1}^K F_k y_{t-k} + G\xi_t \tag{B.4a}$$

$$\mathbb{E}[\xi_t \xi_t'] = I. \tag{B.4b}$$

839 The key idea of the DSGE-VAR estimation proposed in Del Negro et al. (2007) is to estimate

³⁷Although Ireland (2004) does not refer to the residual between model and data as measurement error, other papers in the literature describing his procedure do. Examples are Del Negro and Schorfheide (2009) and Cúrdia and Reis (2012).

840 this time series process with the prior distribution for F and Ω that is centered at the values
841 implied by a DSGE model, $F(\Psi)$ and $G(\Psi)$, where Ψ is the vector containing the parameters
842 of the DSGE model. The estimation procedure consists of jointly estimating Ψ , the structural
843 parameters of the DSGE model, which pin down the prior for the VAR coefficients, and the
844 VAR coefficients themselves.

845 The precision of the prior of the VAR coefficients is controlled with a scalar parameter,
846 λ . If λ is equal to ∞ , then one estimates an unrestricted VAR and if λ is equal to 0, then
847 the procedure boils down to estimating a DSGE without allowing for misspecification. The
848 estimation is executed for different values of λ . To determine the optimal value for λ , the
849 authors propose using the marginal data density, which compares in-sample fit with model
850 complexity. If the restrictions imposed by the DSGE model are incorrect, then the procedure
851 will put more weight on the VAR.

852 As pointed out in Chari et al. (2008), DSGE models often do not imply a VAR represen-
853 tation with a finite number of lags, unless all state variables are included. Thus, not only the
854 DSGE, but also the VAR component of the DSGE-VAR procedure could be misspecified.

855 **Wedges.** Yet another approach to deal with misspecification is to add “wedges” to specific
856 model equations. This procedure was introduced in Chari et al. (2007). Inoue et al. (2015)
857 use this setup to formally test for model misspecification. A wedge may have different inter-
858 pretations or possibly no simple interpretation. From an econometric point a view, wedges
859 are not different from regular structural disturbances in how they affect time series proper-
860 ties of the model. That is, they impose restrictions on the policy functions just as regular
861 structural disturbances do and it matters crucially how one enters wedges. For example,
862 the assumption that a wedge only enters one and not all model equations is a restriction.
863 Although some wedges can enter more than one equation, wedges used in the literature only
864 enter a few specific model equations chosen by the researcher a priori and – as pointed out
865 in Inoue et al. (2015) – wedges can be introduced in many different ways. By contrast, ASDs
866 appear in all equations. If one prefers a more concise specification, then the idea of our ag-
867 nostic procedure indicates involves using a statistical model selection criterion not economic
868 arguments.³⁸

869 Appendix C. Additional discussion for Monte Carlo experiments

870 This appendix starts by giving some additional information about our experiments and
871 continues by providing additional results.

³⁸There may be valid economic or other reasoning to make additional restrictions. For example, in our empirical application we give such motivation for our assumption that the ASDs appear in the flexible-price block of the SW with the same coefficient as in the corresponding equation of the actual model. However, such restrictions make the procedure less agnostic.

872 Appendix C.1. Details

873 We follow SW and do not estimate the depreciation rate, δ , the steady-state wage mark-
874 up, $\bar{\mu}$, the steady-state level of government expenditures, \bar{g} , the curvature in the Kimball
875 goods-market aggregator, ε_p , and the curvature in the Kimball labor-market aggregator, ε_w .
876 Since we use demeaned data, we also fix the trend growth rate, $\bar{\gamma}$, the parameter controlling
877 steady state hours, \bar{l} , the parameter controlling steady state inflation, $\bar{\pi}$, and the discount
878 factor, β .

879 We deviate in one aspect from SW and that is related to the parameter ρ_{ga} , which
880 captures the impact of the TFP structural disturbance on the government expenditures
881 structural disturbance. We set this coefficient equal to zero in both the true *dgp* and in
882 the empirical model. This implies that all structural disturbances are uncorrelated. This
883 is a typical assumption and makes our misspecification experiment more transparent. As
884 discussed below, the misspecification considered is related to the specification of the set of
885 structural disturbances. If $\rho_{ga} \neq 0$, then we would have to make additional choices whenever
886 the misspecification involves either the TFP or the government spending shock. We explored
887 some alternative cases in which $\rho_{ga} \neq 0$ and found similar results.

888 We adopt Maximum Likelihood estimation. This involves a nontrivial optimization given
889 the complexity of the model and number of parameters to be estimated. However, our
890 optimization problem is relatively well defined. The empirical model is very close to the true
891 *dgp*. Moreover, we use the true parameter values as the initial conditions for the optimization
892 routine and we specify bounds for the parameter values. These choices decrease computing
893 time and also give a misspecified model the best possible chance to deliver estimates that
894 are close to the truth. The innovation standard deviations of the disturbances are restricted
895 to be in the interval $[0, 10]$ and the coefficients of their time series process in the interval
896 $[0, 99]$. Given our focus on misspecified disturbances, we want these intervals to be large.
897 For the structural parameters we set the lower bound and the upper bound to the first and
898 ninety-ninth percentile according to the SW prior, centered at the parameter values of the
899 true *dgp*.

900 Although the Monte Carlo experiments focus on ML estimation, we used a Bayesian model
901 comparison for the exercise where we answered whether a researcher would *in practice* reject
902 the wrong empirical specification when compared with the true one. We used a Bayesian ap-
903 proach because Bayesian estimation is the dominant strategy to estimate DSGE models. For
904 this comparison we restricted ourselves to 100 Monte Carlo replication, since each replication
905 involves a computationally intensive MCMC procedure to trace the posterior.

906 Our analysis has some features in common with Ferroni et al. (2015), but there are
907 important differences. They only consider one specific misspecified empirical model whereas
908 we consider a total of forty-two. Although they consider a limited Monte Carlo experiment
909 (with 100 replications), the main discussion focuses on particular sample of 200 observations.
910 Most importantly, their main focus is on the consequences of using an inverse gamma prior
911 for parameters that could well be zero. Our focus is on the misspecification of the empirical
912 model, not the specification of the prior.

913 Appendix C.2. Identification of structural parameters

914 To properly evaluate the misspecification experiments of this paper, it is important that
915 the estimated parameters are identified. If not, then any detected distortions of parameter
916 estimates and implied model properties could be due to lack of identification and not mis-
917 specification. We use the test proposed in Komunjer and Ng (2011) to check whether the
918 parameters of the empirical specifications used in our experiments are identified. We will
919 refer to this test as the KN test. This test provides both necessary and sufficient conditions
920 for local identification under a set of weak conditions.³⁹ It focuses on the state-space repre-
921 sentation of the model and – in contrast to earlier identification tests – does not require the
922 user to specify a set of particular autocovariances.⁴⁰ The results document that parameters
923 are identified in all experiments. In Appendix D.3, we document that weak identification is
924 not an issue either.

925 **Identification of original Smets-Wouters estimation exercise.** SW fix the values
926 of five parameters: depreciation, δ , steady-state wage mark-up, $\bar{\mu}$, steady-state exogenous
927 spending, \bar{g} , curvature in the Kimball goods-market aggregator, ε_p , and curvature in the
928 Kimball labor-market aggregator, ε_w . Komunjer and Ng (2011) consider the identification of
929 the SW model, but their empirical specification is slightly different from the one of SW in
930 that all variables are demeaned. By contrast, the data in the SW estimation exercise does
931 contain information about the level, since the inflation rate and the nominal interest rate are
932 in levels. Komunjer and Ng (2011) show that several subsets of the five parameter restrictions
933 mentioned above are sufficient to obtain identification *if* the parameter controlling steady
934 state hours, \bar{l} , and the parameter controlling steady state inflation, $\bar{\pi}$, are fixed as well. It
935 makes sense that identification requires more restrictions when information about the levels
936 is not used in the estimation.

937 **Identification of our specifications.** The empirical and true specifications used in our
938 Monte Carlo experiments have six structural disturbances, whereas the original SW empirical
939 model has seven. This may imply that less parameters are identified. It is important that
940 the parameters that we try to estimate are identified. If parameters are not identified, then
941 different parameter combinations lead to the same criterion of fit used in the estimation, so it
942 would not be surprising if parameter estimates are different for slightly different specifications.

943 Consequently, we adopt the following conservative strategy to ensure identification. The
944 KN test checks rank conditions of matrices and to see whether there is a singularity one
945 needs to choose a tolerance criterion. We set the criterion at a level that is more strict than
946 the one chosen by KN.⁴¹ We follow SW and fix the values of the five parameters mentioned
947 above. In addition, we fix all parameters that have a direct effect on the means of variables,
948 since we use demeaned variables in the estimation. The associated parameters are the trend

³⁹These are a stability condition and regularity conditions on the innovations.

⁴⁰An example of such an earlier test is Iskrev (2010).

⁴¹We set “Tol” equal to 1e-2 instead of 1e-3 (a higher number means that the test is more difficult to pass).

Table C.5: Komunjer and NG identification test.

| required number | 41 | 225 | 36 | 302 | - | |
|------------------------------|-------------------------------|-------------------------|-------------------------|-----------------------|-------|-----|
| n | $\overline{\Delta}_\Lambda^S$ | $\overline{\Delta}_T^S$ | $\overline{\Delta}_U^S$ | $\overline{\Delta}^S$ | pass? | |
| $\varepsilon_{a,t}$ excluded | 12 | 41 | 225 | 36 | 302 | yes |
| $\varepsilon_{b,t}$ excluded | 12 | 41 | 225 | 36 | 302 | yes |
| $\varepsilon_{g,t}$ excluded | 12 | 41 | 225 | 36 | 302 | yes |
| $\varepsilon_{i,t}$ excluded | 12 | 41 | 225 | 36 | 302 | yes |
| $\varepsilon_{r,t}$ excluded | 12 | 41 | 225 | 36 | 302 | yes |
| $\varepsilon_{w,t}$ excluded | 13 | 41 | 225 | 36 | 302 | yes |
| $\varepsilon_{p,t}$ excluded | 13 | 41 | 225 | 36 | 302 | yes |

Notes. Here, n is the number of restrictions, which includes the number of coefficients fixed in all experiments and the number of coefficients in the law of motion of the excluded exogenous random variable that are all set to zero. $\overline{\Delta}_\Lambda^S$ is a matrix that contains the derivatives of all the vectorized elements in the state-space representation of the model (the A , B , C , D matrices and the variance-covariance matrices) evaluated at the true parameter values. It is intuitive that this matrix needs to have full rank for identification. But it is not sufficient. $\overline{\Delta}_T^S$ and $\overline{\Delta}_U^S$ are matrices with particular elements related to the state-space representation. The matrix $\overline{\Delta}^S = [\overline{\Delta}_\Lambda^S \ \overline{\Delta}_T^S \ \overline{\Delta}_U^S]$ needs to have full rank to pass the KN test.

949 growth rate, $\overline{\gamma}$, the parameter controlling steady state hours, \overline{l} , the parameter controlling
950 steady state inflation, $\overline{\pi}$, and the discount factor, β .⁴² Finally, as discussed in Section 2.1,
951 we fix the spillover from the productivity disturbance to exogenous spending and set it equal
952 to zero.

953 The results of the KN test are reported in Table C.5 and it indicates that the identification
954 test is passed in all cases. That is, lack of identification is not driving the results in this paper.

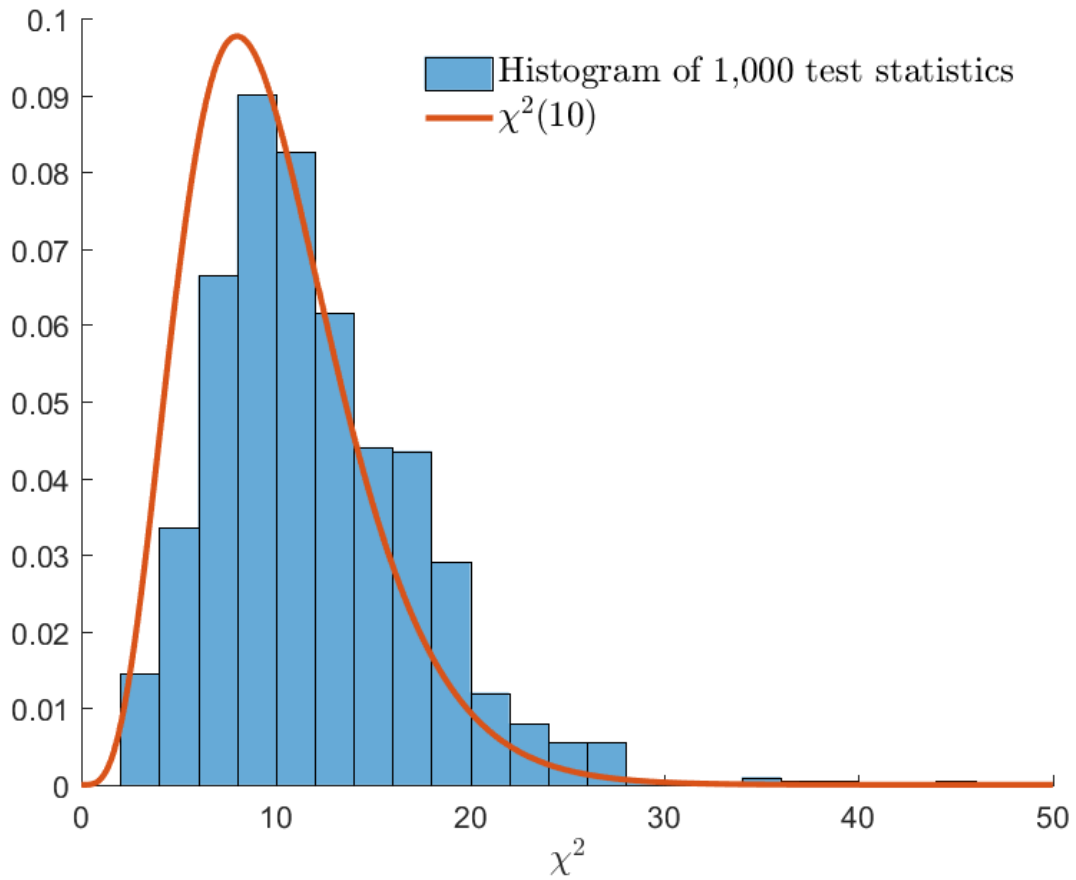
955 Appendix C.3. Additional results

956 Figures C.6 and C.7 plot the histograms of the estimated χ^2 statistics across Monte Carlo
957 replications for the two experiments of Section 3 together with the theoretical (large-sample)
958 χ^2 distribution. The number of degrees of freedom is equal to 10. They document that the
959 distribution of test statistics across Monte Carlo replications is quite close to the theoretical
960 one.

961 Tables C.6 and C.7 document detailed information on the distribution of parameter esti-
962 mates for the two Monte Carlo experiments.

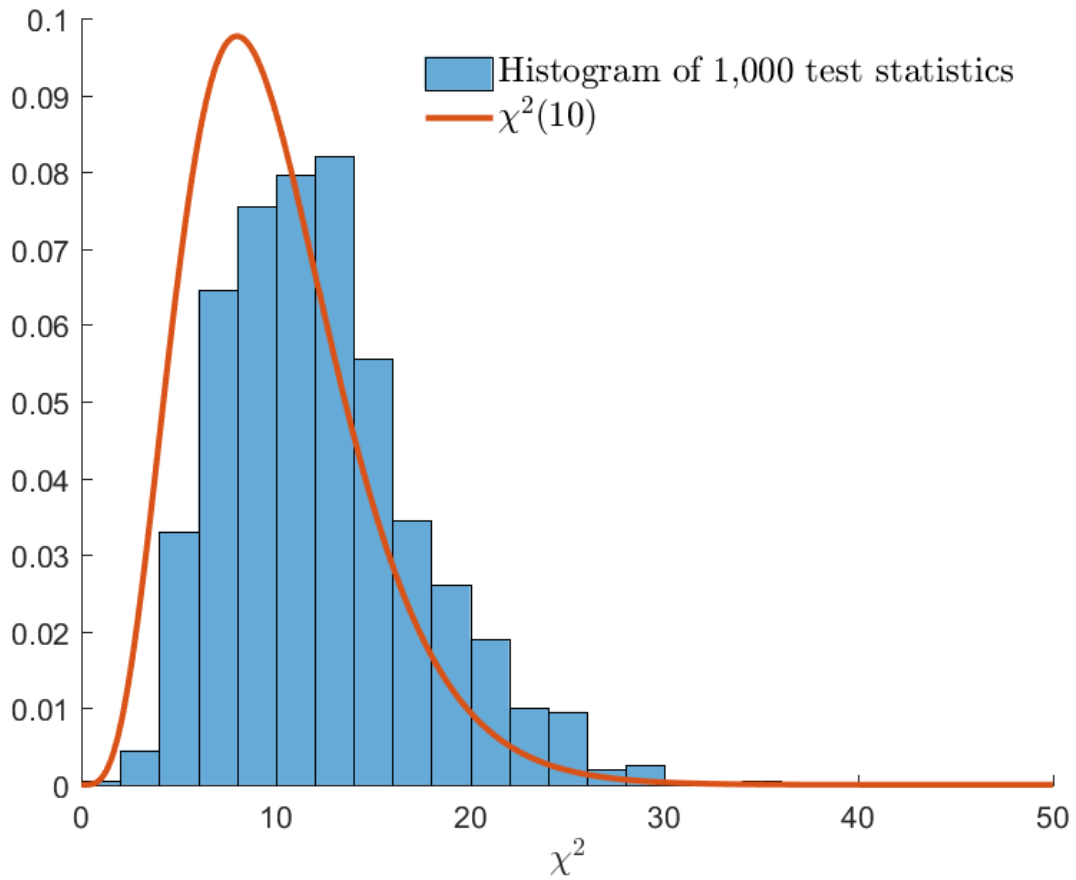
⁴²It is a conservative choice to fix all four, since identification only requires that two parameters are fixed according to the test of Komunjer and Ng (2011).

Figure C.6: Likelihood ratio test agnostic versus fully-specified model: First experiment



Notes. The figure plots the distribution of χ^2 statistics of the first Monte Carlo experiment and the theoretical distribution according to large sample theory. This Monte Carlo experiment corresponds to the case when the true *dgp* does not include a monetary policy disturbance, but the empirical model leaves out the investment disturbance instead.

Figure C.7: Likelihood ratio test agnostic versus fully-specified model: Second experiment



Notes. The figure plots the distribution of χ^2 statistics of the first Monte Carlo experiment and the theoretical distribution according to large sample theory. This Monte Carlo experiment corresponds to the case when the true *dgp* does not include a TFP disturbance, but the empirical model leaves out the investment disturbance instead.

Table C.6: Parameter estimates across Monte Carlo replications: First experiment

| | Truth | LB | UB | misspecified estimation | | | | | ASD procedure | | | | | SW specification | | | | |
|----------------|-------|-------|-------|-------------------------|------|------|------|------|---------------|------|------|------|------|------------------|------|------|------|------|
| | | | | 10% | 25% | 50% | 75% | 90% | 10% | 25% | 50% | 75% | 90% | 10% | 25% | 50% | 75% | 90% |
| α | 0.19 | 0.07 | 0.31 | 0.07 | 0.07 | 0.08 | 0.10 | 0.14 | 0.15 | 0.18 | 0.21 | 0.23 | 0.26 | 0.15 | 0.17 | 0.20 | 0.22 | 0.24 |
| σ_c | 1.39 | 0.53 | 2.25 | 1.35 | 1.44 | 1.66 | 2.16 | 2.24 | 1.16 | 1.29 | 1.47 | 1.69 | 1.95 | 1.19 | 1.29 | 1.42 | 1.58 | 1.81 |
| Φ | 1.61 | 1.33 | 1.89 | 1.46 | 1.69 | 1.86 | 1.89 | 1.89 | 1.33 | 1.34 | 1.57 | 1.85 | 1.89 | 1.33 | 1.38 | 1.60 | 1.83 | 1.89 |
| ϕ | 5.48 | 1.99 | 8.97 | 3.42 | 5.03 | 6.48 | 8.05 | 8.87 | 3.18 | 3.78 | 4.59 | 5.63 | 6.74 | 3.80 | 4.27 | 4.91 | 5.55 | 6.24 |
| λ | 0.71 | 0.45 | 0.90 | 0.62 | 0.71 | 0.79 | 0.84 | 0.86 | 0.57 | 0.62 | 0.66 | 0.71 | 0.76 | 0.61 | 0.64 | 0.68 | 0.72 | 0.76 |
| ξ_w | 0.73 | 0.47 | 0.92 | 0.58 | 0.65 | 0.71 | 0.78 | 0.86 | 0.57 | 0.62 | 0.68 | 0.75 | 0.82 | 0.58 | 0.63 | 0.69 | 0.75 | 0.81 |
| σ_ℓ | 1.92 | 0.18 | 3.66 | 0.18 | 0.18 | 0.18 | 0.19 | 0.54 | 0.89 | 1.16 | 1.55 | 2.17 | 2.77 | 1.07 | 1.36 | 1.76 | 2.24 | 2.71 |
| ξ_p | 0.65 | 0.40 | 0.86 | 0.69 | 0.74 | 0.79 | 0.83 | 0.85 | 0.50 | 0.56 | 0.63 | 0.69 | 0.73 | 0.52 | 0.57 | 0.63 | 0.68 | 0.73 |
| ι_w | 0.59 | 0.24 | 0.89 | 0.24 | 0.28 | 0.38 | 0.53 | 0.68 | 0.32 | 0.44 | 0.60 | 0.76 | 0.88 | 0.34 | 0.47 | 0.61 | 0.76 | 0.88 |
| ι_p | 0.22 | 0.01 | 0.65 | 0.04 | 0.13 | 0.25 | 0.38 | 0.51 | 0.02 | 0.08 | 0.16 | 0.25 | 0.34 | 0.03 | 0.08 | 0.16 | 0.24 | 0.31 |
| ψ | 0.54 | 0.20 | 0.86 | 0.26 | 0.38 | 0.59 | 0.75 | 0.84 | 0.38 | 0.47 | 0.57 | 0.69 | 0.81 | 0.41 | 0.48 | 0.57 | 0.66 | 0.77 |
| r_π | 2.03 | 1.45 | 2.61 | 1.58 | 1.78 | 2.05 | 2.33 | 2.55 | 1.71 | 1.88 | 2.10 | 2.41 | 2.60 | 1.76 | 1.91 | 2.09 | 2.35 | 2.57 |
| ρ | 0.81 | 0.53 | 0.97 | 0.74 | 0.77 | 0.80 | 0.82 | 0.83 | 0.78 | 0.80 | 0.82 | 0.84 | 0.85 | 0.78 | 0.80 | 0.82 | 0.84 | 0.85 |
| r_y | 0.08 | -0.04 | 0.20 | 0.05 | 0.07 | 0.11 | 0.17 | 0.20 | 0.05 | 0.07 | 0.08 | 0.11 | 0.13 | 0.06 | 0.07 | 0.08 | 0.10 | 0.12 |
| $r_{\Delta y}$ | 0.22 | 0.10 | 0.34 | 0.11 | 0.15 | 0.17 | 0.18 | 0.19 | 0.19 | 0.21 | 0.22 | 0.23 | 0.24 | 0.20 | 0.21 | 0.22 | 0.23 | 0.24 |
| ρ_a | 0.95 | 0.00 | 0.99 | 0.60 | 0.93 | 0.95 | 0.96 | 0.96 | 0.88 | 0.92 | 0.94 | 0.95 | 0.96 | 0.90 | 0.93 | 0.94 | 0.96 | 0.96 |
| ρ_b | 0.18 | 0.00 | 0.99 | 0.03 | 0.08 | 0.16 | 0.27 | 0.75 | 0.03 | 0.08 | 0.14 | 0.20 | 0.26 | 0.04 | 0.09 | 0.15 | 0.21 | 0.26 |
| ρ_g | 0.97 | 0.00 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.94 | 0.96 | 0.97 | 0.98 | 0.98 | 0.94 | 0.96 | 0.97 | 0.98 | 0.98 |
| ρ_p | 0.90 | 0.00 | 0.99 | 0.50 | 0.66 | 0.78 | 0.91 | 0.97 | 0.68 | 0.80 | 0.87 | 0.92 | 0.95 | 0.74 | 0.82 | 0.88 | 0.92 | 0.94 |
| ρ_w | 0.97 | 0.00 | 0.99 | 0.93 | 0.97 | 0.99 | 0.99 | 0.99 | 0.93 | 0.95 | 0.97 | 0.98 | 0.99 | 0.94 | 0.96 | 0.97 | 0.98 | 0.99 |
| μ_p | 0.74 | 0.00 | 0.99 | 0.21 | 0.38 | 0.64 | 0.86 | 0.94 | 0.31 | 0.48 | 0.62 | 0.73 | 0.80 | 0.38 | 0.51 | 0.64 | 0.73 | 0.80 |
| μ_w | 0.88 | 0.00 | 0.99 | 0.72 | 0.81 | 0.87 | 0.91 | 0.94 | 0.73 | 0.80 | 0.85 | 0.89 | 0.92 | 0.76 | 0.82 | 0.86 | 0.89 | 0.92 |
| σ_a | 0.45 | 0.00 | 10.00 | 0.62 | 0.70 | 0.85 | 1.05 | 1.22 | 0.35 | 0.38 | 0.42 | 0.48 | 0.53 | 0.37 | 0.39 | 0.44 | 0.48 | 0.53 |
| σ_b | 0.24 | 0.00 | 10.00 | 0.06 | 0.20 | 0.24 | 0.26 | 0.28 | 0.21 | 0.22 | 0.24 | 0.26 | 0.28 | 0.21 | 0.23 | 0.24 | 0.26 | 0.28 |
| σ_g | 0.52 | 0.00 | 10.00 | 0.51 | 0.53 | 0.55 | 0.57 | 0.59 | 0.48 | 0.49 | 0.52 | 0.54 | 0.56 | 0.48 | 0.50 | 0.52 | 0.54 | 0.56 |
| σ_p | 0.14 | 0.00 | 10.00 | 0.12 | 0.14 | 0.15 | 0.17 | 0.18 | 0.11 | 0.12 | 0.14 | 0.15 | 0.17 | 0.11 | 0.12 | 0.14 | 0.15 | 0.16 |
| σ_w | 0.24 | 0.00 | 10.00 | 0.19 | 0.20 | 0.22 | 0.24 | 0.25 | 0.21 | 0.23 | 0.24 | 0.26 | 0.28 | 0.21 | 0.23 | 0.25 | 0.26 | 0.28 |

Notes. The table provides information on the distribution of the indicated parameter across the Monte Carlo replications. See Table 1 for the definitions of the parameters. This Monte Carlo experiment corresponds to the case when the true dgp does not include a monetary policy disturbance, but the empirical model leaves out the investment disturbance instead.

Table C.7: Parameter estimates across Monte Carlo replications: Second experiment

| | Truth | LB | UB | misspecified estimation | | | | | ASD procedure | | | | | SW specification | | | | |
|----------------|-------|-------|-------|-------------------------|------|------|------|------|---------------|------|------|------|------|------------------|------|------|------|------|
| | | | | 10% | 25% | 50% | 75% | 90% | 10% | 25% | 50% | 75% | 90% | 10% | 25% | 50% | 75% | 90% |
| α | 0.19 | 0.07 | 0.31 | 0.09 | 0.12 | 0.16 | 0.21 | 0.26 | 0.14 | 0.18 | 0.21 | 0.25 | 0.29 | 0.14 | 0.17 | 0.20 | 0.23 | 0.26 |
| σ_c | 1.39 | 0.53 | 2.25 | 1.39 | 1.57 | 2.07 | 2.24 | 2.25 | 1.14 | 1.27 | 1.45 | 1.75 | 2.14 | 1.16 | 1.26 | 1.41 | 1.63 | 1.93 |
| Φ | 1.61 | 1.33 | 1.89 | 1.39 | 1.58 | 1.79 | 1.87 | 1.89 | 1.33 | 1.35 | 1.58 | 1.85 | 1.89 | 1.34 | 1.43 | 1.61 | 1.78 | 1.87 |
| ϕ | 5.48 | 1.99 | 8.97 | 4.09 | 4.99 | 6.23 | 7.43 | 8.48 | 3.29 | 3.85 | 4.57 | 5.43 | 6.54 | 3.85 | 4.34 | 4.98 | 5.64 | 6.44 |
| λ | 0.71 | 0.45 | 0.90 | 0.54 | 0.61 | 0.69 | 0.76 | 0.80 | 0.55 | 0.61 | 0.66 | 0.72 | 0.77 | 0.60 | 0.64 | 0.68 | 0.73 | 0.77 |
| ξ_w | 0.73 | 0.47 | 0.92 | 0.55 | 0.61 | 0.67 | 0.72 | 0.78 | 0.55 | 0.60 | 0.67 | 0.74 | 0.81 | 0.58 | 0.62 | 0.69 | 0.75 | 0.80 |
| σ_ℓ | 1.92 | 0.18 | 3.66 | 0.18 | 0.18 | 0.20 | 0.55 | 1.03 | 0.46 | 0.89 | 1.56 | 2.58 | 3.54 | 0.67 | 1.07 | 1.78 | 2.64 | 3.38 |
| ξ_p | 0.65 | 0.40 | 0.86 | 0.62 | 0.68 | 0.73 | 0.78 | 0.83 | 0.49 | 0.54 | 0.62 | 0.68 | 0.73 | 0.51 | 0.56 | 0.62 | 0.68 | 0.72 |
| ι_w | 0.59 | 0.24 | 0.89 | 0.24 | 0.30 | 0.42 | 0.57 | 0.69 | 0.32 | 0.46 | 0.61 | 0.78 | 0.89 | 0.34 | 0.47 | 0.62 | 0.77 | 0.88 |
| ι_p | 0.22 | 0.01 | 0.65 | 0.04 | 0.11 | 0.24 | 0.34 | 0.46 | 0.01 | 0.07 | 0.15 | 0.24 | 0.32 | 0.03 | 0.08 | 0.16 | 0.24 | 0.32 |
| ψ | 0.54 | 0.20 | 0.86 | 0.30 | 0.40 | 0.53 | 0.67 | 0.80 | 0.37 | 0.45 | 0.57 | 0.70 | 0.81 | 0.40 | 0.48 | 0.58 | 0.68 | 0.79 |
| r_π | 2.03 | 1.45 | 2.61 | 1.74 | 2.04 | 2.32 | 2.54 | 2.60 | 1.59 | 1.86 | 2.28 | 2.59 | 2.61 | 1.68 | 1.89 | 2.22 | 2.52 | 2.60 |
| ρ | 0.81 | 0.53 | 0.97 | 0.79 | 0.81 | 0.84 | 0.86 | 0.87 | 0.73 | 0.77 | 0.81 | 0.84 | 0.86 | 0.76 | 0.78 | 0.81 | 0.84 | 0.85 |
| r_y | 0.08 | -0.04 | 0.20 | 0.08 | 0.10 | 0.13 | 0.16 | 0.19 | 0.04 | 0.06 | 0.10 | 0.12 | 0.14 | 0.05 | 0.07 | 0.09 | 0.12 | 0.13 |
| $r_{\Delta y}$ | 0.22 | 0.10 | 0.34 | 0.11 | 0.13 | 0.17 | 0.20 | 0.22 | 0.17 | 0.19 | 0.23 | 0.26 | 0.29 | 0.17 | 0.20 | 0.22 | 0.25 | 0.28 |
| ρ_b | 0.18 | 0.00 | 0.99 | 0.07 | 0.14 | 0.22 | 0.33 | 0.53 | 0.02 | 0.08 | 0.14 | 0.21 | 0.26 | 0.04 | 0.09 | 0.15 | 0.21 | 0.26 |
| ρ_g | 0.97 | 0.00 | 0.99 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.94 | 0.96 | 0.97 | 0.98 | 0.98 | 0.94 | 0.96 | 0.97 | 0.98 | 0.98 |
| ρ_r | 0.12 | 0.00 | 0.99 | 0.00 | 0.00 | 0.02 | 0.06 | 0.10 | 0.00 | 0.04 | 0.11 | 0.18 | 0.23 | 0.01 | 0.05 | 0.11 | 0.17 | 0.22 |
| ρ_p | 0.90 | 0.00 | 0.99 | 0.48 | 0.65 | 0.78 | 0.91 | 0.97 | 0.74 | 0.82 | 0.89 | 0.92 | 0.95 | 0.78 | 0.84 | 0.89 | 0.92 | 0.95 |
| ρ_w | 0.97 | 0.00 | 0.99 | 0.94 | 0.96 | 0.97 | 0.98 | 0.99 | 0.93 | 0.95 | 0.97 | 0.98 | 0.99 | 0.94 | 0.96 | 0.97 | 0.98 | 0.98 |
| μ_p | 0.74 | 0.00 | 0.99 | 0.17 | 0.37 | 0.61 | 0.84 | 0.92 | 0.36 | 0.50 | 0.62 | 0.72 | 0.81 | 0.44 | 0.55 | 0.65 | 0.73 | 0.80 |
| μ_w | 0.88 | 0.00 | 0.99 | 0.73 | 0.79 | 0.84 | 0.89 | 0.91 | 0.73 | 0.79 | 0.84 | 0.89 | 0.91 | 0.76 | 0.81 | 0.85 | 0.89 | 0.92 |
| σ_b | 0.24 | 0.00 | 10.00 | 0.14 | 0.19 | 0.22 | 0.24 | 0.26 | 0.20 | 0.22 | 0.24 | 0.26 | 0.28 | 0.21 | 0.23 | 0.24 | 0.26 | 0.28 |
| σ_g | 0.52 | 0.00 | 10.00 | 0.49 | 0.51 | 0.53 | 0.55 | 0.57 | 0.47 | 0.49 | 0.51 | 0.54 | 0.56 | 0.48 | 0.50 | 0.52 | 0.54 | 0.56 |
| σ_r | 0.24 | 0.00 | 10.00 | 0.22 | 0.23 | 0.24 | 0.25 | 0.26 | 0.22 | 0.23 | 0.24 | 0.25 | 0.26 | 0.22 | 0.23 | 0.24 | 0.25 | 0.26 |
| σ_p | 0.14 | 0.00 | 10.00 | 0.12 | 0.14 | 0.15 | 0.16 | 0.18 | 0.11 | 0.12 | 0.14 | 0.15 | 0.17 | 0.11 | 0.12 | 0.14 | 0.15 | 0.16 |
| σ_w | 0.24 | 0.00 | 10.00 | 0.20 | 0.21 | 0.23 | 0.24 | 0.26 | 0.21 | 0.23 | 0.25 | 0.27 | 0.28 | 0.22 | 0.23 | 0.25 | 0.27 | 0.28 |

Notes. The table provides information on the distribution of the indicated parameter across the Monte Carlo replications. See Table 1 for the definitions of the parameters. This Monte Carlo experiment corresponds to the case when the true dgp does not include a TFP disturbance, but the empirical model leaves out the investment disturbance instead.

963 Appendix D. Large sample consequences of misspecification

964 ASDs are designed to deal with the misspecification of structural disturbances. In Section
965 3, we used Monte Carlo experiments to document that the consequences of such misspecifi-
966 cation for parameter estimates can be quite severe.

967 The analysis of that section has some drawbacks. First, the results are subject to sampling
968 variation, since the sample size was chosen to resemble the length of data series available
969 to macroeconomists in practice. Consequently, the documented deviations from the truth
970 may not be due to misspecification solely but also to small sample issues such as bias and,
971 or course, sampling uncertainty. Second, the analysis only focused on the consequences
972 for parameter estimates whereas it also would be interesting to look at model properties
973 as implied by parameter estimates. Examples are impulse response functions (IRFs) and
974 moments of model variables. Third, we only discussed the results for two representative
975 experiments, whereas there are forty-two possible experiments.

976 In this appendix, we study the consequences of misspecification in greater detail. First,
977 by using samples of 100,000 observations we reduce sampling variation to negligible levels.
978 Thus, all deviations from the true values are due to misspecification. Second, in addition to
979 parameter estimates we also look at implied moments of model variables and implied IRFs.
980 Third, we consider all possible forty-two experiments. In all other aspects, the experiment is
981 identical to the one described in Section 3.

982 Appendix D.1. Consequences for parameter values

983 Table D.8 reports some key percentiles (across experiments) to characterize the range of
984 the estimated parameter values. When constructing percentiles, we only consider parameters
985 that are in both the true and empirical specification.⁴³ All parameter estimates are affected
986 by misspecification to some extent. Moreover, the minor misspecifications considered in these
987 forty-two experiments lead to massive distortions for several parameters.

988 The median parameter estimates (across experiments) are relatively close to the true pa-
989 rameter values. Thus, our choice of experiments does not favor bias in a particular direction.
990 There is one exception. The median value of the estimated standard deviation of the pro-
991 ductivity disturbance innovation, σ_a , is equal to 0.92 compared to a true value of 0.45. The
992 reason is that this disturbance often “absorbs” the variation of the disturbance that is not
993 included in the empirical specification. Thus, the disturbance that is wrongly included in the
994 empirical specification does not necessarily fulfill this role.

995 Even if we exclude cases for which the estimates fall in the bottom or top 10%, then we
996 find that estimates are substantially different from their true value for many parameters. For
997 example, for the labor supply elasticity with respect to the real wage, σ_l , the 10th percentile
998 is equal to 0.18 and the 90th percentile is equal to 3.66, compared with a true value of

⁴³Specifically, for the parameters of the exogenous random processes, the experiments in which the dis-
turbance is part of the empirical model – but not part of the true *dgp* – are excluded from the calculations
of the percentiles.

Table D.8: Parameter values: Point estimates across misspecification experiments

| | Truth | Imposed Min | Min | 10% | 25% | Median | 75% | 90% | Max | Imposed Max |
|----------------|--------------|------------------------|------------|------------|------------|---------------|------------|------------|------------|------------------------|
| α | 0.19 | 0.07 | 0.07 | 0.11 | 0.17 | 0.19 | 0.20 | 0.23 | 0.31 | 0.31 |
| σ_c | 1.39 | 0.53 | 0.53 | 0.78 | 1.14 | 1.35 | 1.60 | 1.82 | 2.25 | 2.25 |
| Φ | 1.61 | 1.33 | 1.33 | 1.33 | 1.53 | 1.77 | 1.89 | 1.89 | 1.89 | 1.89 |
| ϕ | 5.48 | 1.99 | 2.71 | 3.59 | 5.47 | 7.38 | 8.97 | 8.97 | 8.97 | 8.97 |
| λ | 0.71 | 0.45 | 0.45 | 0.59 | 0.71 | 0.74 | 0.84 | 0.89 | 0.90 | 0.90 |
| ξ_w | 0.73 | 0.47 | 0.50 | 0.67 | 0.73 | 0.75 | 0.82 | 0.87 | 0.91 | 0.92 |
| σ_ℓ | 1.92 | 0.18 | 0.18 | 0.18 | 0.52 | 1.87 | 2.71 | 3.66 | 3.66 | 3.66 |
| ξ_p | 0.65 | 0.40 | 0.53 | 0.60 | 0.65 | 0.78 | 0.86 | 0.86 | 0.86 | 0.86 |
| ι_w | 0.59 | 0.24 | 0.24 | 0.27 | 0.38 | 0.58 | 0.61 | 0.80 | 0.89 | 0.89 |
| ι_p | 0.22 | 0.01 | 0.01 | 0.01 | 0.10 | 0.22 | 0.32 | 0.48 | 0.63 | 0.65 |
| ψ | 0.54 | 0.20 | 0.20 | 0.20 | 0.42 | 0.54 | 0.68 | 0.86 | 0.86 | 0.86 |
| r_π | 2.03 | 1.45 | 1.45 | 1.45 | 1.71 | 2.07 | 2.39 | 2.61 | 2.61 | 2.61 |
| ρ | 0.81 | 0.53 | 0.62 | 0.73 | 0.79 | 0.81 | 0.85 | 0.88 | 0.92 | 0.97 |
| r_y | 0.08 | -0.04 | -0.04 | 0.01 | 0.05 | 0.09 | 0.16 | 0.20 | 0.20 | 0.20 |
| $r_{\Delta y}$ | 0.22 | 0.10 | 0.10 | 0.10 | 0.10 | 0.20 | 0.24 | 0.34 | 0.34 | 0.34 |
| ρ_a | 0.95 | 0.00 | 0.50 | 0.82 | 0.92 | 0.96 | 0.98 | 0.99 | 0.99 | 0.99 |
| ρ_b | 0.18 | 0.00 | 0.04 | 0.09 | 0.13 | 0.17 | 0.26 | 0.36 | 0.80 | 0.99 |
| ρ_g | 0.97 | 0.00 | 0.94 | 0.96 | 0.97 | 0.97 | 0.99 | 0.99 | 0.99 | 0.99 |
| ρ_I | 0.71 | 0.00 | 0.57 | 0.60 | 0.68 | 0.71 | 0.78 | 0.84 | 0.95 | 0.99 |
| ρ_r | 0.12 | 0.00 | 0.01 | 0.06 | 0.11 | 0.13 | 0.18 | 0.33 | 0.50 | 0.99 |
| ρ_p | 0.90 | 0.00 | 0.70 | 0.77 | 0.84 | 0.89 | 0.93 | 0.96 | 0.98 | 0.99 |
| ρ_w | 0.97 | 0.00 | 0.93 | 0.95 | 0.97 | 0.97 | 0.98 | 0.99 | 0.99 | 0.99 |
| μ_p | 0.74 | 0.00 | 0.08 | 0.22 | 0.43 | 0.73 | 0.82 | 0.91 | 0.95 | 0.99 |
| μ_w | 0.88 | 0.00 | 0.00 | 0.00 | 0.87 | 0.89 | 0.92 | 0.96 | 0.98 | 0.99 |
| σ_a | 0.45 | 0.00 | 0.42 | 0.47 | 0.67 | 0.92 | 1.49 | 2.57 | 3.20 | 10 |
| σ_b | 0.24 | 0.00 | 0.07 | 0.20 | 0.23 | 0.24 | 0.26 | 0.27 | 0.29 | 10 |
| σ_g | 0.52 | 0.00 | 0.52 | 0.52 | 0.52 | 0.53 | 0.55 | 0.56 | 0.57 | 10 |
| σ_I | 0.45 | 0.00 | 0.14 | 0.25 | 0.39 | 0.44 | 0.46 | 0.48 | 0.54 | 10 |
| σ_r | 0.24 | 0.00 | 0.22 | 0.23 | 0.23 | 0.24 | 0.26 | 0.28 | 0.31 | 10 |
| σ_p | 0.14 | 0.00 | 0.04 | 0.09 | 0.12 | 0.14 | 0.15 | 0.16 | 0.17 | 10 |
| σ_w | 0.24 | 0.00 | 0.18 | 0.20 | 0.21 | 0.24 | 0.25 | 0.29 | 0.31 | 10 |

Notes. This table gives information about the parameter estimates across the forty-two misspecification experiments. For the parameters of the laws of motion of the disturbances, we exclude an experiment from the calculations of the percentiles when the disturbance is part of the empirical model, but not part of the true *dgp*. The table also reports the bounds imposed on parameter estimates. See Table 1 for the definitions of the parameters.

999 1.92. For the parameter capturing the indexation of wages ι_w , the same two percentiles
1000 are 0.27 and 0.80, compared with a true value of 0.59. For the parameter capturing the
1001 indexation of prices, ι_p , the two numbers are 0.01 and 0.48, compared with a true value of
1002 0.22. When the two 10% tails are not excluded and the full range of estimates is considered,
1003 then the range substantially increases. Specifically, the largest values are 0.89 and 0.63 for
1004 the indexation of wages and prices, respectively.⁴⁴ Recall that these distortions are solely
1005 due to misspecification, not to small-sample variation.

1006 For several parameters, the results remain bad when we narrow the range of outcomes
1007 considered. For example, when we exclude the bottom and the top 25%, then the values for
1008 σ_l , vary between 0.52 and 2.71 compared with a true value of 1.92. The results are also quite
1009 bad for ϕ , the elasticity in the capital adjustment cost function, for which the 25th percentile
1010 is equal to 5.47 and the 75th percentile is equal to 8.97.

1011 Appendix D.2. Consequences for model properties

1012 The previous section documents that misspecification can lead to large distortions in
1013 parameter values. Parameter estimates are often of interest in themselves. At least as
1014 important are the properties of the estimated structural model. It could be that different
1015 parameter configurations lead to similar model properties. In this section, we address this
1016 by looking at implied moments and IRFs.

1017 Appendix D.2.1. Implied model moments

1018 We begin by documenting the consequences of model misspecification for implied model
1019 moments using the misspecification setup described above. Table D.9 reports the range of
1020 values for typical business cycle properties as implied by the estimated parameter values of the
1021 forty-two experiments considered. Specifically, it reports standard deviations and correlation
1022 coefficients relative to their true values. Thus, a value equal to 1 means that there is no
1023 distortion. The column labeled “true value” reports the range of values the corresponding
1024 moment has according to the true dgp .⁴⁵

1025 Misspecification implies an upward bias for volatility in our experiments.⁴⁶ This upward
1026 bias could be specific to our particular type of misspecification. However, the observed
1027 upward bias is consistent with the simple analytical example discussed in Appendix D.5.⁴⁷

⁴⁴Parameter estimates are constrained to be in a range, and the largest estimate of the wage indexation parameter is constrained by the imposed upper bound.

⁴⁵Moments are not the same across experiments, since we adjust the standard deviations of the structural disturbances to ensure that the wrongly omitted disturbance does not play an important role.

⁴⁶Section Appendix D.1 documents an upward bias for σ_a , the standard deviation of the TFP disturbance. Since one disturbance is missing from the empirical model, it is not surprising that there is a shift towards some of the other disturbances. By contrast, here we find an upward bias for *total* variability.

⁴⁷In Appendix D.5, we discuss a simple example which documents analytically how maximum likelihood estimation of a misspecified model can lead to an *arbitrarily* large upward bias in the implied variance of an observable.

Table D.9: Moments: Ratio of implied value to truth across experiments with misspecification

| | True value (across experiments) | Min | 10% | 25% | Median | 75% | 90% | Max |
|----------------------|------------------------------------|-----------------------------------|------|------|--------|------|------|-------|
| | | (estimates, scaled by true value) | | | | | | |
| Std(y_t) | [3.48 , 5.12] | 0.51 | 0.78 | 0.92 | 1.03 | 1.64 | 4.46 | 6.03 |
| Std(c_t) | [3.30 , 5.58] | 0.45 | 0.76 | 0.92 | 1.03 | 1.81 | 4.12 | 6.62 |
| Std(i_t) | [9.73 , 12.94] | 0.70 | 0.87 | 0.99 | 1.11 | 1.71 | 3.81 | 6.47 |
| Std(r_t) | [0.52 , 0.61] | 0.76 | 0.90 | 0.94 | 1.00 | 1.36 | 2.28 | 2.78 |
| Std(π_t) | [0.37 , 0.54] | 0.64 | 0.72 | 0.94 | 1.01 | 1.25 | 2.21 | 2.98 |
| Std(w_t) | [2.13 , 2.70] | 0.73 | 0.83 | 0.92 | 1.08 | 2.28 | 5.57 | 10.87 |
| Corr(y_t, c_t) | [0.65 , 0.94] | 0.28 | 0.68 | 0.93 | 0.99 | 1.07 | 1.15 | 1.52 |
| Corr(y_t, i_t) | [0.74 , 0.87] | 0.69 | 0.83 | 0.95 | 1.00 | 1.10 | 1.16 | 1.29 |
| Corr(c_t, i_t) | [0.63 , 0.89] | -0.68 | 0.60 | 0.92 | 1.00 | 1.19 | 1.34 | 1.57 |
| Corr(c_t, r_t) | [-0.65 , -0.35] | -0.71 | 0.54 | 0.86 | 0.99 | 1.11 | 1.52 | 2.13 |
| Corr(i_t, w_t) | [0.29 , 0.69] | -1.52 | 0.10 | 0.64 | 1.07 | 1.49 | 1.99 | 3.28 |
| Corr(i_t, π_w) | [0.51 , 0.80] | 0.36 | 0.84 | 0.97 | 1.02 | 1.17 | 1.34 | 1.75 |

Notes. This table reports the outcomes across experiments for the indicated moment as implied by parameter estimates relative to its true value. Thus a value equal to 1 indicates that there is no distortion due to misspecification. Each row reports percentiles across our forty-two experiments. It also reports the range of values of the true moments across the experiments. All moments considered are related to variables that are used in the estimation as observables.

1028 The results are solely due to misspecification, since we use very large samples and our ML
1029 estimator is consistent when the empirical model is correctly specified.

1030 The overestimation of volatility is enormous in some cases. Even if we exclude the top
1031 25%, then standard deviations can be multiples of the true standard deviation. For example,
1032 the 75th percentile for the standard deviation of wages is 2.28 times its true value. This ratio
1033 increases to 5.57 when we only exclude the top 10%. The 90th percentiles for the consumption
1034 and output standard deviation ratios are 4.12 and 4.46, which also indicates massive over-
1035 prediction. The 90th percentile for investment is equal to 3.81 and in the worst experiment the
1036 implied standard deviation is 6.47 times as big as the true value. By contrast, the values in
1037 the lower tail are less drastic. Excluding the bottom 10%, we find that the largest distortions
1038 are found for inflation for which the 10th percentile is 0.72, that is, implied volatility is 28%
1039 below its true value. If we consider all experiments, then the smallest ratio is equal to 0.45,
1040 which is found for the implied standard deviation of consumption.

1041 Misspecification also has large quantitative implications for correlation coefficients. In
1042 fact, the sign of the correlation coefficient as implied by parameter estimates turns out to be
1043 different from its sample analogue in several cases. This would not be a big deal if the two
1044 correlation coefficients are both close to zero. But there are also cases in which the implied
1045 correlation coefficient according to the estimated empirical model and the true correlation

1046 coefficient are both large in absolute value and differ in sign.^{48,49}

1047 Appendix D.2.2. Impulse response functions (IRFs)

1048 To conclude the discussion on the consequences of misspecification, we document that
1049 misspecification can also have a large impact on impulse response functions. There are many
1050 IRFs to consider. Figure D.8 plots for three IRFs the outcomes across the experiments and
1051 documents that the distortions can be large. We exclude the cases when the disturbance
1052 of interest is in the empirical specification, but not part of the true *dgp*. It would not be
1053 surprising if these are different.⁵⁰ Thus, the disturbance of interest is part of the true *dgp* as
1054 well as the empirical model for all three cases considered.

1055 Figure D.8a plots the response of output to a TFP disturbance. This is obviously a key
1056 characteristic of the model. The black line plots the true IRF and the grey lines plot the
1057 IRFs as implied by the empirical model for the different experiments. All IRFs are based the
1058 same size shock.⁵¹ If the grey lines are close to the black line, then misspecification of the
1059 empirical model has only minor consequences for the IRF considered. The sign of the IRF
1060 is virtually always correct and TFP disturbances always have a noticeable positive impact
1061 on aggregate output.⁵² Nevertheless, the figure documents that there are large differences in
1062 terms of initial impact, overall magnitude, shape, and persistence.

1063 Figure D.8b plots the response of the real wage to a monetary policy shock. This is clearly
1064 the kind of model property one would want to get right when analyzing monetary policy. The
1065 figure shows again a wide variety of responses across the different empirical specifications.
1066 Whereas the true response is substantial, there are several empirical specifications that pre-
1067 dict a very small change. There are also a few specifications that give a much larger response.
1068 We want to reemphasize that the plotted IRFs are for a disturbance that is correctly included
1069 in the empirical model.

1070 Figure D.8c reports the results for the inflation IRF of an investment-specific shock. For
1071 most experiments the IRFs display a similar pattern, but there are important differences in
1072 terms of magnitude. For three experiments, however, the IRFs are completely at odds with
1073 the true IRF. Whereas the true IRF is positive and has reverted back to zero after twenty
1074 periods, the IRFs implied by these three misspecified empirical models are negative and
1075 indicate larger volatility and more persistence. Again, relatively small changes in parameter

⁴⁸A striking example is the experiment in which the government disturbance is not present in the true *dgp* and the empirical model excludes the risk-premium disturbance instead. The true correlation between consumption and investment is equal to 0.67 whereas the one implied by the estimated model is equal to -0.41.

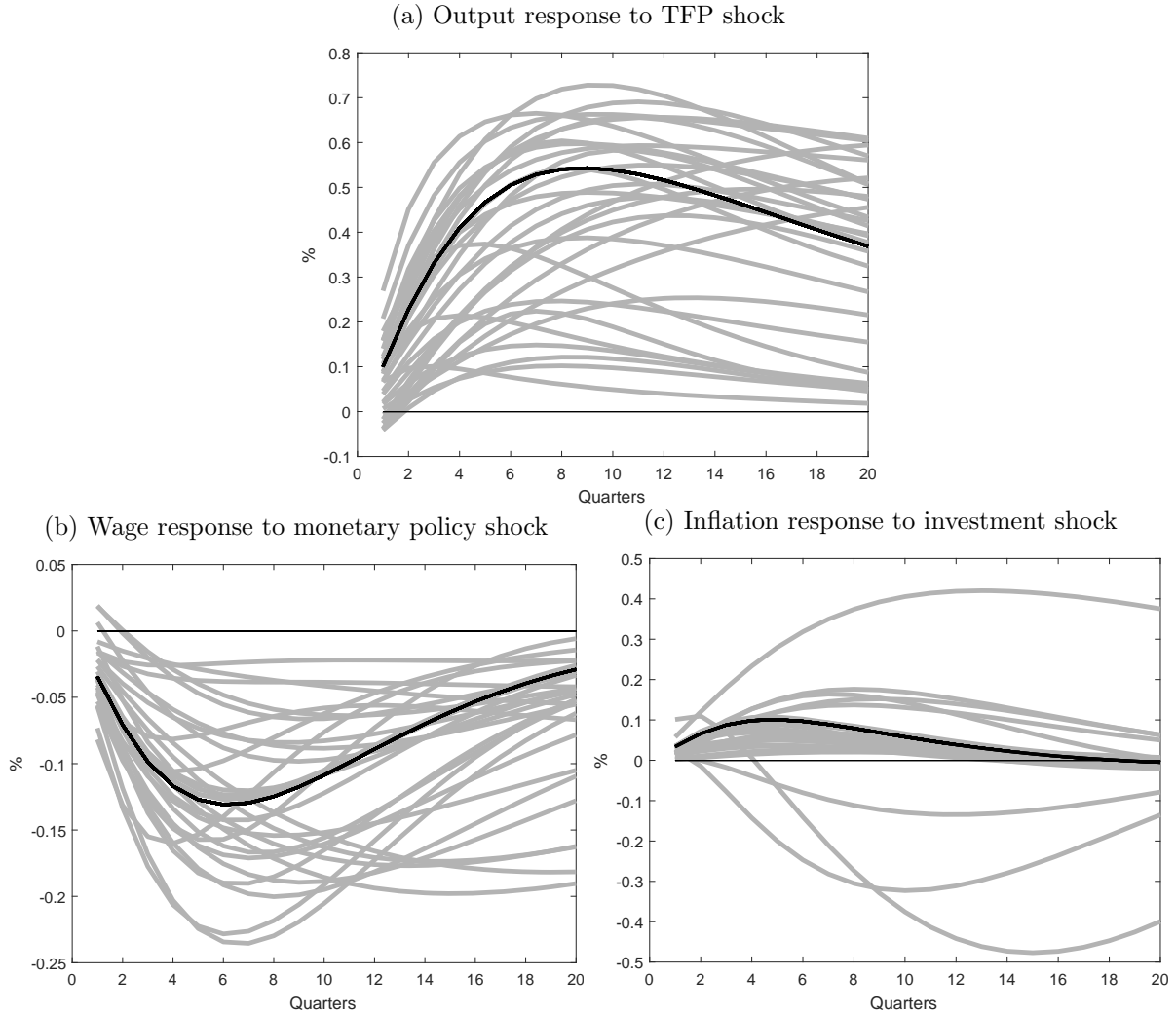
⁴⁹The smallest correlation coefficient (in absolute value) according to the true model is 0.29, so any sign change implies a nontrivial change in the correlation coefficient.

⁵⁰Also, we cannot calculate IRFs for a particular disturbance if that disturbance is not part of the empirical specification. This means that each figure plots IRFs for thirty-two cases.

⁵¹That is, one standard deviation according to the original SW model. Differences across IRFs are bigger if we use the estimated standard deviations for the different experiments.

⁵²In some experiments, the initial response is negative. However, its value is then very small.

Figure D.8: IRFs according to true (black) and misspecified (grey) empirical models



Notes. The figure plots the true IRF (black) and the IRFs implied by the misspecified (grey) empirical models considered. The results are based on a very large sample, so results are not due to small sample variation. These IRFs are for shocks that are correctly included in the model. Also, we do not use estimated standard deviations, but use the same size shock for all IRFs.

1076 values can change these IRFs such that they are much closer to the true IRF.⁵³

⁵³Specifically, if σ_c , the parameter controlling curvature in the utility function and λ , the parameter indicating the habit component in the utility function, are set equal to their true values, then these three IRFs have a shape that is similar to the true IRF, that is, also predict a hump-shaped positive response. The responses still differ somewhat from the truth in having a more delayed response and a more persistent effect. The estimated values for σ_c in the three experiments are 0.65, 0.53, and 0.53, whereas the true value is equal to 1.39. The estimated values for λ are equal to 0.86, 0.87, and 0.85, whereas the true value is equal to 0.71.

1077 **Appendix D.3. Is weak identification the cause?**

1078 In Appendix C.2, it was shown that all parameters are identified in all models consid-
1079 ered.⁵⁴ Moreover, we use a very large sample to estimate the parameters so the large range of
1080 values for parameter estimates cannot be caused by samples being too short to be informa-
1081 tive. Also, the finding that the different parameter values are associated with quite different
1082 model properties indicates that the results discussed in this section are not due to param-
1083 eters not being identified. As a final check, we compare the values of the likelihood function
1084 according to the misspecified model at the estimated values and the true values. When using
1085 the true values, we do re-estimate the parameters of the exogenous random variables.⁵⁵ The
1086 smallest difference between the two log likelihood values is equal to 14.5 and there are only
1087 four experiments for which the difference is less than 100. The mean (median) difference is
1088 equal to 10,371 (5501).⁵⁶

1089 **Appendix D.4. Choosing Monte Carlo experiments**

1090 A careful Monte Carlo experiment requires a sufficiently large number of replications. We
1091 use 1,000. Each replication involves a computationally intensive optimization routine. This
1092 means we would not be able to do a small-sample version of all 42 experiments in this ap-
1093 pendix. The two we use in section 3 were chosen as follows. We ranked all experiments by the
1094 likelihood value obtained with the misspecified specification relative to the likelihood value
1095 obtained with the correct specification. The idea is that misspecification is less severe if the
1096 difference in likelihood values is smaller. The first experiment chosen is the one correspond-
1097 ing to the sixty-sixth percentile and the second is the one corresponding to the thirty-third
1098 percentile.⁵⁷ Thus, our experiments are neither the least nor the most problematic in terms
1099 of misspecification.

1100 **Appendix D.5. An analytical example**

1101 In this section, we give a *very* simple example to indicate that misspecification can have
1102 large distortive effects in the sense that *implied* properties of the model using the parameter
1103 estimates can be at odds with the *actual* corresponding properties of the data that are used to

⁵⁴All true specifications have one structural disturbance less than the original SW model. This turns out not to matter for identification. In fact, estimated parameters remain identified when we do the identification test for specifications with five disturbances that exclude the disturbance that is not part of the true *dgp* as well as the one that is erroneously omitted from the empirical specification.

⁵⁵This is a conservative choice, since differences in the likelihoods would be larger if these parameters are not re-estimated.

⁵⁶It is not surprising that across experiments, there are some for which the misspecification is smaller than for others resulting in smaller differences between the two likelihood values. After all, our experiments are not designed to find large misspecification. Our set is constructed using a simple variation in the set of the original structural disturbances.

⁵⁷The first (second) Monte Carlo experiment corresponds to the case when the true *dgp* does not include a monetary policy (TFP) disturbance, but the empirical model leaves out the investment disturbance instead.

1104 estimate the parameters. The model is linear, and all variables have a Normal distribution.
 1105 Throughout this section, parameter estimates are based on population moments. Thus,
 1106 the results are not due to small sample variation. The estimation procedure is Maximum
 1107 Likelihood (ML).

1108 More specifically, this example demonstrates that there can be massive differences between
 1109 the variances of observables as *implied* by the model using estimated parameter values and
 1110 the actual variances in the data set. This result is surprising since the ML estimator of
 1111 the variance of a given time series is the sample variance when the variable has a Normal
 1112 distribution. We will show that this is not necessarily true for implied variances when the
 1113 empirical model is misspecified.⁵⁸

1114 **True model.** The true model is given by the following set of equations:

$$y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} = \Lambda \varepsilon_t, \quad (\text{D.1})$$

$$\mathbb{E} [\varepsilon_t \varepsilon_t'] = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}, \quad (\text{D.2})$$

1115 and we make the following assumption about the distribution of the error terms:

$$\varepsilon_{1,t} \sim N(0, \sigma_1^2) \text{ and } \varepsilon_{2,t} \sim N(0, \sigma_2^2). \quad (\text{D.3})$$

1116 **Misspecification.** The objective is to estimate the standard deviations of the structural
 1117 disturbances, σ_1^2 and σ_2^2 . The researcher takes the value of Λ as given. The empirical model
 1118 is misspecified, because $\bar{\Lambda} \neq \Lambda$ is used instead of the true value.

1119 **Empirical specifications.** We consider the following two empirical specifications:

Case 1: Empirical model given by

$$y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \bar{\Lambda} \varepsilon_t, \quad \mathbb{E} [\varepsilon_t \varepsilon_t'] = \begin{bmatrix} \bar{\sigma}_1^2 & \bar{\sigma}_{12} \\ \bar{\sigma}_{12} & \bar{\sigma}_2^2 \end{bmatrix}. \quad (\text{D.4})$$

⁵⁸As a byproduct of this paper, we learned that there also can be large gaps between *actual* properties of the data used and the corresponding *implied* properties according to the Maximum Likelihood estimates of the model parameters when the DSGE model is correctly specified, but a data sample with finite length is used. Since the objective of Maximum Likelihood is not to match moments, there is no reason why there should be a close match, but we were surprised by the large magnitudes of the differences. For example, using a sample of 1,000 observations generated by the SW model with seven disturbances and the correct empirical specification, it is not unusual to find implied standard deviations for the observables that are three to five times their data counterpart. Such differences will disappear as the sample size increases, since the estimator is consistent, but such asymptotic results do not provide much assurance if there is a small sample bias even at a relatively large sample size of 1,000 observations.

Case 2: Empirical model given by

$$y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \bar{\Lambda} \varepsilon_t, \quad \mathbb{E} [\varepsilon_t \varepsilon_t'] = \begin{bmatrix} \bar{\sigma}_1^2 & 0 \\ 0 & \bar{\sigma}_2^2 \end{bmatrix}. \quad (\text{D.5})$$

1120 Both empirical models are misspecified, because they use the wrong value of Λ . In the
 1121 first case, the empirical model allows the correlation between the two innovations to be non-
 1122 zero even though it is equal to zero according to the true data generating process. In the
 1123 second case, the empirical model imposes that the correlation is equal to zero, just as it is in
 1124 the true model.

1125 **Case 1: Wrong Λ and allow for wrong σ_{12} .** Since the model is linear and the shocks
 1126 have a normal distribution, the ML estimator of the variance-covariance matrix $\mathbb{E} [\varepsilon_t \varepsilon_t']$, $\hat{\Sigma}_\varepsilon$,
 1127 is given by

$$\hat{\Sigma}_\varepsilon = \bar{\Lambda}^{-1} \hat{\Sigma}'_y \bar{\Lambda}^{-1'}. \quad (\text{D.6})$$

1128 As mentioned above, we abstract from sampling variation and $\hat{\Sigma}'_y$ is estimated using popula-
 1129 tion moments. This means that the ML estimator of $\hat{\Sigma}'_\varepsilon$ is given by

$$\hat{\Sigma}_\varepsilon = \bar{\Lambda}^{-1} \mathbb{E} [y_t y_t'] \bar{\Lambda}^{-1'} \quad (\text{D.7})$$

$$= \bar{\Lambda}^{-1} \Lambda \Lambda' \bar{\Lambda}^{-1'}. \quad (\text{D.8})$$

True versus implied variance. The purpose of this section is to document the conse-
 quences of misspecification for the implied variance of the observable y_t according to the
 estimated model. The *true* variance-covariance matrix is given by:

$$\Sigma_y^{\text{true}} = \mathbb{E} [y_t y_t'] = \Lambda \Lambda'. \quad (\text{D.9})$$

1130 The *implied* variance of y_t according the researcher's (misspecified) model, $\hat{\Sigma}_y$, is given by

$$\hat{\Sigma}_y = \bar{\Lambda} \hat{\Sigma}_\varepsilon \bar{\Lambda}' \quad (\text{D.10})$$

$$= \bar{\Lambda} \bar{\Lambda}^{-1} \Lambda \Lambda' \bar{\Lambda}^{-1'} \bar{\Lambda}' \quad (\text{D.11})$$

$$= \Lambda \Lambda' = \Sigma_y^{\text{true}}. \quad (\text{D.12})$$

1131 Thus, the procedure actually generates the correct answer even though an incorrect empirical
 1132 specification is used. In this case, the estimated empirical model is misspecified for two
 1133 reasons, namely it has the wrong Λ and the estimated value of σ_{12} is not equal to its true value.
 1134 These have exactly offsetting effects in terms of their impact on the implied variance. Another
 1135 way to look at this result is the following. By allowing for a more flexible specification, i.e.,
 1136 a non-zero value for σ_{12} , the researcher would get a better answer for the implied variance of
 1137 y_t even though the flexibility implies that the estimated model is wrong in more dimensions.

Case 2: Wrong Λ and correct σ_{12} . Obtaining the estimate for $\widehat{\Sigma}_\varepsilon$ is just as easy as in the previous case. Given $\widehat{\Lambda}$ and data for y_t , one can calculate the values for ε_t and use these to calculate the variance of ε_t and the implied variance of y_t . The following is a complicated, but useful way to express the outcome:

$$\widehat{\Sigma}_\varepsilon = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \overline{\Lambda}^{-1} \Lambda \Lambda' \overline{\Lambda}^{-1'} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \overline{\Lambda}^{-1} \Lambda \Lambda' \overline{\Lambda}^{-1'} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \quad (\text{D.13})$$

1138 **True versus implied variance.** The implied variance of y_t is equal to

$$\widehat{\Sigma}_y = \left(\begin{array}{c} \overline{\Lambda} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \overline{\Lambda}^{-1} \Lambda \Lambda' \overline{\Lambda}^{-1'} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \overline{\Lambda}' \\ + \\ \overline{\Lambda} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \overline{\Lambda}^{-1} \Lambda \Lambda' \overline{\Lambda}^{-1'} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \overline{\Lambda}' \end{array} \right) \neq \Lambda \Lambda' = \Sigma_y^{\text{true}} \quad (\text{D.14})$$

The reason $\widehat{\Sigma}_y \neq \Sigma_y^{\text{true}}$ is that the $\overline{\Lambda}$ terms do not cancel out. In our Monte Carlo experiments with misspecified models, we find that there often are large gaps between the variances of the observables used in the estimation and the corresponding variances as implied by the model using the estimated parameters. Moreover, there is a bias. That is, the implied variance is typically larger than the actual variance. Our Monte Carlo experiments are a lot more complicated than this example, but this example may shed light on the coincidence of high implied variances. Specifically, because the $\overline{\Lambda}$ s do not cancel out, the expression for $\widehat{\Sigma}_y$ contains terms like the following:

$$\overline{\Lambda} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \overline{\Lambda}^{-1} = \frac{1}{\overline{\lambda}_{11}\overline{\lambda}_{22} - \overline{\lambda}_{12}\overline{\lambda}_{21}} \begin{bmatrix} \overline{\lambda}_{11}\overline{\lambda}_{22} & -\overline{\lambda}_{11}\overline{\lambda}_{12} \\ \overline{\lambda}_{21}\overline{\lambda}_{22} & -\overline{\lambda}_{12}\overline{\lambda}_{21} \end{bmatrix}. \quad (\text{D.15})$$

1139 This equation documents that the ratio of the implied variance relative to the true variance
 1140 could be arbitrarily large if the term in the denominator goes to zero.⁵⁹ For a correctly
 1141 specified model this would not matter, since the small term in the denominator would then
 1142 be offset by an equally small term in the numerator. But this is not necessarily the case for
 1143 an incorrectly specified model.

⁵⁹The opposite is less likely, since it would require values for the λ_{ij} coefficients such that the combinations appearing in square brackets are small, but the particular combination in the denominator is not. For example, one cannot accomplish this by simply choosing small values for the λ_{ij} terms.

1144 Appendix E. ASD procedure for the Smets-Wouters model

1145 In this appendix, we provide further details on how the ASD procedure is implemented
1146 in Section 4 and we provide additional results.

1147 Appendix E.1. Including ASDs in SW equations

1148 To apply the ASD procedure to the SW model, we adapt the Dynare program provided
1149 by the authors.⁶⁰ Adapting a Dynare program to add an agnostic disturbance is easy. Specif-
1150 ically, for the first ASD, $\tilde{\varepsilon}_{A,t}$, we do the following.

- 1151 1. In the model block, we add $\tilde{\Upsilon}_{j,A}\tilde{\varepsilon}_{A,t}$ to the j^{th} equation, where $\tilde{\varepsilon}_{A,t}$ is the agnostic
1152 disturbance and $\tilde{\Upsilon}_{j,A}$ the coefficient associated with the agnostic disturbance in the j^{th}
1153 equation. Details are given below.⁶¹
- 1154 2. We add an equation to the model block that describes the law of motion for $\tilde{\varepsilon}_{A,t}$. If the
1155 agnostic disturbance replaces a regular structural disturbance, then this disturbance
1156 should be taken out of the program.
- 1157 3. Declare $\tilde{\varepsilon}_{A,t}$ as a variable and declare the elements of $\tilde{\Upsilon}_{j,A}$ and the coefficients of the
1158 law of motion for $\tilde{\varepsilon}_{A,t}$ as parameters.
- 1159 4. Specify a prior for the elements of $\tilde{\Upsilon}_{j,A}$.

1160 We do not add the agnostic disturbance to Equations (6) and (12) of the SW model,
1161 because these equations just contain definitions for capacity utilization and the wage mark-
1162 up, respectively.⁶² The set of equations for the SW model consists of two parts. The first
1163 part models the flexible price economy and the second part models the actual economy with
1164 sticky prices. One needs to model the flexible-price economy, because the flexible-price output
1165 level is used to define the output gap, which is one of the arguments in the monetary policy
1166 rule. In principle, one could let the agnostic disturbance enter the equations of the sticky-
1167 price economy and the associated equations in the flexible-price economy with a different
1168 coefficient.⁶³ Given the minor role played by the flexible-price block, it doesn't quite make
1169 sense to introduce so many additional parameters. Moreover, structural disturbances would
1170 enter the associated pair of equations in the same way in most economic models. Therefore,
1171 we also restrict the agnostic disturbance to enter the associated equations in the same way.

⁶⁰The program is available at <https://www.aeaweb.org/articles?id=10.1257/aer.97.3.586> under the "Download Data Set" link.

⁶¹The other two ASDs are added using the same procedure.

⁶²Equation numbers refer to those in Smets and Wouters (2007). We do allow the agnostic disturbances to affect the utilization rate and the wage mark-up directly by including it in the model equations that specify their relationship with other model variables.

⁶³The sticky-price block contains some equations, such as the monetary policy rule, that do not have a counterpart in the flexible-price economy.

1172 The exception is SW Equation (13) because it captures both potential stickiness in wages
1173 and the relationship between the wage rate and its mark-up.

1174 Specifically, we add the agnostic disturbance to Equations (1), (2), (3), (4), (5), (7), (8),
1175 (9), and (11) of the SW model and the associated equations of the flexible-price economy. We
1176 also add it to Equation (13) in both the flexible and the sticky-price part of the model, but
1177 here we allow coefficients to differ. In addition, we add the agnostic disturbance to Equations
1178 (10) and (14) which do not have a counterpart in the flexible-price economy. This means
1179 that $\tilde{\Upsilon}_{.,A}$ has thirteen elements. The last coefficient associated with the agnostic disturbance
1180 is the autoregressive coefficient of its law of motion. The standard deviation of the agnostic
1181 disturbance is normalized to be equal to 1.

1182 **Additional information.** The SW specification uses consumption growth as an observable
1183 and has an equation that defines consumption growth. Allowing an agnostic disturbance
1184 to affect this equation would capture measurement error (which would be correlated with
1185 structural disturbances if this ASD also appears in other model equations with a non-zero
1186 coefficient). We do not explore this possibility to keep the analysis parsimonious and to stay
1187 close the SW approach, which does not allow for measurement error.

1188 As pointed out in the main text, the prior mean of $\tilde{\varepsilon}_{A,t}$ and $\tilde{\varepsilon}_{B,t}$ are set equal to associated
1189 values of $\varepsilon_{b,t}$ and $\varepsilon_{i,t}$. For example, suppose we use the ASD procedure to test the restrictions
1190 of the risk-premium disturbance by replacing it with an ASD. The risk-premium disturbance
1191 appears in two equations, namely the consumption/bond Euler equation and the capital-
1192 valuation equation. The prior means of the reduced-form agnostic coefficients for these
1193 two equations are set equal to the values according to the SW restrictions with structural
1194 parameters evaluated at their prior means. The reduced-form coefficients associated with
1195 the other equations have a prior mean equal to zero. Having a non-zero prior has a practical
1196 advantage. The signs of the coefficients of an agnostic disturbance are not identified. That is,
1197 one can switch the signs of the coefficients of an ASD as long as one does it for all coefficients.
1198 A necessary consequence of its agnostic nature is that the sign of an ASD has no a priori
1199 meaning. If the prior means of all ASD coefficients are zero, then the ASD coefficients can
1200 flip sign for different runs of the MCMC procedure.

1201 Appendix E.2. Model selection procedures

1202 **Which structural disturbances to include?** The first stage of the model selection pro-
1203 cedure is to decide which regular and agnostic structural disturbances to include. Specifically,
1204 we compare a set of models that do or do not include the risk-premium disturbance, that do
1205 or do not include the investment disturbance, and that include one, two, or three ASDs.⁶⁴ We
1206 still allow the risk-premium and the investment-specific disturbance to appear in the final set
1207 even though replacement by an ASD improved model fit. The reason is that a specification

⁶⁴To estimate the model with all seven observables, an empirical specification with only one ASD would need either the risk-premium or the investment disturbance to avoid a singularity.

1208 with both ASDs and these regular disturbances could perform even better.

Table E.10: Model selection procedure for SW model: Step 1

| regular structural | | agnostic | | | marginal |
|---------------------|---------------------|-----------------------------|-----------------------------|-----------------------------|----------------|
| $\varepsilon_{b,t}$ | $\varepsilon_{i,t}$ | $\tilde{\varepsilon}_{A,t}$ | $\tilde{\varepsilon}_{B,t}$ | $\tilde{\varepsilon}_{C,t}$ | data density |
| no | no | yes | yes | no | -906.85 |
| no | no | yes | yes | yes | -925.55 |
| no | yes | yes | no | no | -908.79 |
| no | yes | yes | yes | no | -907.46 |
| no | yes | yes | yes | yes | -922.94 |
| yes | no | no | yes | no | -919.81 |
| yes | no | yes | yes | no | -907.32 |
| yes | no | yes | yes | yes | -921.71 |
| yes | yes | no | no | no | -922.40 |
| yes | yes | yes | no | no | -909.35 |
| yes | yes | no | yes | no | -920.26 |
| yes | yes | yes | yes | no | -908.09 |
| yes | yes | yes | yes | yes | -922.82 |

Notes. The table reports the marginal data density for different empirical specifications regarding three agnostic disturbances and the two disturbances that are misspecified, that is, the risk-premium disturbance, $\varepsilon_{b,t}$, and the investment disturbance, $\varepsilon_{i,t}$. The number in bold indicates the highest outcome.

1209 Table E.10 reports the results. It shows that the model with the highest marginal data
 1210 density is one with two agnostic disturbances, without the SW risk-premium, and without
 1211 the SW investment-specific disturbance. Another indication that there is no need for these
 1212 two SW structural disturbances is that their role in terms of explaining variation in the data
 1213 is very small when agnostic disturbances are included. According to the (unconditional)
 1214 variance decomposition of the estimated SW model, the risk-premium disturbance is espe-
 1215 cially important for the price of capital, consumption growth, and output growth explaining
 1216 45.4%, 61.2%, and 22.1% of total variability, respectively. It only plays a minor role for other
 1217 variables. When agnostic disturbances are added, then these three numbers drop to 3.88%,
 1218 3.88%, and 2.05%, respectively.⁶⁵ The reduction in the role of the investment disturbance is
 1219 even stronger. In the SW model, the investment disturbance plays a quantitatively important
 1220 role for many variables. For investment growth it even explains 82.1% of the volatility. With
 1221 agnostic disturbances added, its role becomes minuscule. Even for investment growth it only
 1222 explains 0.31%.

1223 **Obtaining a concise ASD specification.** To interpret ASDs, we could use the best
 1224 specification found so far. However, interpretation of an ASD is easier when the specification
 1225 is more concise. To determine whether an agnostic disturbance should be excluded from

⁶⁵These numbers are based on the specification with two ASDs and all seven SW structural disturbances using posterior mode estimates.

1226 some equations, we implement model selection procedures using the marginal data density
1227 as the criterion of fit. This statistic increases when fit improves, but also penalizes additional
1228 parameters.

1229 We consider both a specific-to-general procedure and a general-to-specific procedure and
1230 we apply the procedure for the specifications with two and three ASDs.⁶⁶ The details of these
1231 procedures are described further below. The specific-to-general procedure with three ASDs
1232 leads to the highest MDD and the selected outcome will be our preferred empirical model. The
1233 specific-to-general procedure with two ASDs and the general-to-specific procedure with two
1234 ASDs lead to slightly lower MDDs.⁶⁷ Moreover, the models selected by these three procedures
1235 are very similar. Specifically, the additional ASD in the specification with three ASDs only
1236 plays a minor role. The zero restrictions imposed for the other two ASDs are not exactly the
1237 same, but the differences are due to coefficients that turn out to be small. As documented in
1238 Appendix E.3, the estimates of the parameters are similar and the estimates obtained with
1239 these three empirical specifications imply similar model properties. The general-to-specific
1240 procedure with three ASDs leads to a specification that has a much lower MDD.⁶⁸

1241 In our preferred specification, the first agnostic disturbance enters eight of the thirteen
1242 equations, the second in three, and the third in five. By contrast, the original SW risk-
1243 premium and the investment-specific disturbance appear in only two.

1244 **Details of the model selection procedures.** The general-to-specific model selection
1245 procedure starts with the specification in which the agnostic disturbances are allowed to
1246 enter each model equation. It then calculates the marginal data densities for all possible
1247 specifications in which the ASD is *not* allowed to enter *one* of the model equations. Thus, we
1248 estimate a set of models, each having one less coefficient. If none of the specifications lead
1249 to a better fit, then the procedure stops. If improvements are found, then the procedure is
1250 repeated using the specification that led to the biggest improvement as the benchmark.

1251 The specific-to-general procedure starts with the specifications in which each of the two
1252 ASDs are allowed to enter only one model equation. To avoid a singularity, one cannot

⁶⁶An informal alternative selection procedure would be the following. One starts at the same point as the general-to-specific procedure, that is, with ASDs included in every equation. The marginal posteriors of the agnostic coefficients provides information on the lack of importance of different agnostic coefficients and may provide the researcher promising combinations of zero restrictions to impose. In fact, the posteriors for the coefficients with the fully unrestricted ASD specifications are very predictive of the equations selected by the specific-to-general procedures for this application. Of course, there are good reasons why this informal procedure is not a generally accepted model selection procedure and we cannot expect this to always work well.

⁶⁷The specific-to-general procedure generates an MDD equal to -892.92 with two ASDs and -890.76 with three. The general-to-specific with two ASDs results in an MDD of -894.94 .

⁶⁸Namely, -909.48 . The general-to-specific procedure already stops after two steps. That is, the procedure does not detect that imposing *multiple* restrictions *simultaneously* does lead to substantial improvements. One has to impose some structure on any model selection procedure, because it would be impossible to consider all possible combinations. That is, one has to give instructions on what paths to follow and which ones to ignore. But this means that the model selection procedure may not find the best model. This motivates our use of different model selection criteria.

1253 start with a more parsimonious model.⁶⁹ In the next step, we estimate a set of models in
1254 which one of the ASDs is added to one equation and, thus, one additional parameter is
1255 estimated. The procedure stops if none of the specifications leads to an improvement. If
1256 there is an improvement, then the specification with the largest improvement becomes the
1257 next benchmark and the procedure is repeated.

1258 **Why not consider even more general specifications?** Although our model selection
1259 procedures consider a rich set of models, they are not the most general. Unfortunately, there
1260 are practical limitations to what is feasible. Five SW disturbance are always included in
1261 our specifications. The most ideal setup would be flexible in this dimension as well and not
1262 safeguard any of the seven SW regular disturbances and allow for the possibility of including
1263 seven ASDs (or more). With such a setup all SW disturbances could be replaced by an ASD.
1264 The first problem one would have to deal with is that identification of structural parameters
1265 is likely to limit the number of regular structural disturbances one can replace with ASDs.
1266 Let us consider a simple setup in which there are seven equations for seven state variables
1267 and all state variables are observables. Moreover, each equation has one regular structural
1268 disturbance. A general-to-specific procedure would be complicated since the first-stage model
1269 would have a large number of coefficients to estimate. Specifically, if all seven ASDs appear in
1270 all equations, then one needs to estimate forty-nine reduced-form coefficients. One may need
1271 a rich data set to identify all of them. In our application, the number of coefficients would be
1272 equal to ninety-one, since we have thirteen equations. The specific-to-general procedure faces
1273 the problem that each specification needs at least seven disturbances to avoid singularities.
1274 This means that there are a large number of different models one can start with. For the
1275 simple setup with seven equations described above, this would mean that there are already
1276 $2^7 = 128$ different models to consider in the first round alone.

1277 **Different prior for ASD coefficients.** When we narrow the prior of the agnostic coeffi-
1278 cients by reducing the standard deviation to 0.1, then the restrictions of the monetary policy
1279 disturbance are also rejected. But the increase in the marginal data density is relatively
1280 small, namely from -922.40 to -920.82. The less informed prior of the main text is more
1281 consistent with the idea of the ASDs being agnostic disturbances.

1282 **Appendix E.3. Additional results**

1283 **Correlation of the estimated innovations.** Tables E.11 and E.12 report the contem-
1284 poraneous correlation coefficients of the estimated innovations for the ASD and SW specifi-
1285 cation, respectively. We use the posterior mean estimates to construct the smoothed shocks.
1286 For the SW specification with seven innovations, nine correlation coefficients are significantly

⁶⁹The posteriors of the ASD coefficients in the fully agnostic model provide clear evidence that one of the ASDs is very important for the bond Euler equation and one for the investment Euler equation. So these are natural choices.

1287 different from zero at the 10% or lower level. For the ASD specification with eight innova-
 1288 tions, four coefficients are significant and only two when we exclude the eighth innovation
 1289 associated with $\tilde{\varepsilon}_{C,t}$.

Table E.11: Cross-correlation of innovations: SW

| | η_a | η_g | η_r | η_p | η_w | η_b | η_i |
|----------|----------|----------|----------|----------|----------|----------|----------|
| η_a | . | 0.010 | -0.024 | -0.135* | 0.150 | 0.116 | -0.260** |
| η_g | . | . | 0.166** | 0.218** | -0.160** | -0.262** | -0.074 |
| η_r | . | . | . | -0.056 | -0.048 | 0.186* | 0.037 |
| η_p | . | . | . | . | -0.098 | -0.221** | 0.070 |
| η_w | . | . | . | . | . | -0.019 | -0.189** |
| η_b | . | . | . | . | . | . | 0.152 |

Notes. * (**) indicates significant at the 10% (5%) level. Standard errors are calculated using the VARHAC estimator of Den Haan and Levin (1997) which corrects for serial correlation.

Table E.12: Cross-correlation of innovations: ASD

| | $\tilde{\eta}_a$ | $\tilde{\eta}_g$ | $\tilde{\eta}_r$ | $\tilde{\eta}_p$ | $\tilde{\eta}_w$ | $\tilde{\eta}_A$ | $\tilde{\eta}_B$ | $\tilde{\eta}_C$ |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| $\tilde{\eta}_a$ | . | 0.027 | 0.010 | 0.006 | 0.103 | 0.009 | -0.088 | 0.066 |
| $\tilde{\eta}_g$ | . | . | 0.150 | 0.073 | 0.189** | 0.041 | -0.049 | -0.116 |
| $\tilde{\eta}_r$ | . | . | . | -0.054 | -0.014 | -0.029 | -0.043 | -0.047 |
| $\tilde{\eta}_p$ | . | . | . | . | 0.260** | -0.021 | 0.011 | -0.176** |
| $\tilde{\eta}_w$ | . | . | . | . | . | -0.051 | 0.038 | 0.629** |
| $\tilde{\eta}_A$ | . | . | . | . | . | . | 0.062 | 0.003 |
| $\tilde{\eta}_B$ | . | . | . | . | . | . | . | -0.153 |

Notes. * (**) indicates significant at the 10% (5%) level. Standard errors are calculated using the VARHAC estimator of Den Haan and Levin (1997) which corrects for serial correlation.

Table E.13: Auto-correlation of innovations

| | ASD | | SW |
|------------------|----------|----------|----------|
| $\tilde{\eta}_a$ | -0.060 | η_a | -0.040 |
| $\tilde{\eta}_g$ | -0.024 | η_g | -0.182** |
| $\tilde{\eta}_r$ | -0.170 | η_r | -0.013 |
| $\tilde{\eta}_p$ | -0.121 | η_p | -0.077* |
| $\tilde{\eta}_w$ | 0.069 | η_w | -0.043 |
| $\tilde{\eta}_A$ | -0.069 | η_b | -0.071* |
| $\tilde{\eta}_B$ | -0.155** | η_i | -0.148** |
| $\tilde{\eta}_C$ | -0.245** | | |

Notes. * (**) indicates significant at the 10% (5%) level. Standard errors are calculated using the VARHAC estimator of Den Haan and Levin (1997) which corrects for serial correlation.

1290 Table E.13 reports the auto-correlation coefficients for both empirical specifications.
 1291 Again the ASD specification does quite a bit better with only two significant coefficients

1292 (at the 10% level) for its eight innovations compared to four of the seven for the SW specifi-
1293 cation.

1294 **Impact on parameter estimates and model properties.** Table E.14 documents there
1295 are several differences between the estimated values of the structural parameters obtained
1296 with the fully structural SW specification and our preferred agnostic specification with three
1297 ASDs. For example, the inflation coefficient in the Taylor rule is equal to 2.05 in the SW
1298 specification and 1.77 in ours.⁷⁰ The SW estimate is right at the upper bound of our 90%
1299 highest posterior density (HPD) interval. The SW mean estimate for the parameter charac-
1300 terizing the share of fixed cost in production is equal to 1.61 which is quite a bit higher than
1301 our mean estimate of 1.47 and outside our 90% HPD interval. Also, the mean posterior value
1302 of the MA coefficient of the wage mark-up disturbance is equal to 0.85 according to the SW
1303 specification and 0.59 according to ours. Our mean estimate for the standard deviation of
1304 this disturbance is roughly a third of the SW estimate.

1305 Although there are some nontrivial differences, they are relatively small and the IRFs
1306 of the five regular structural disturbances that are included in both specifications are very
1307 similar for the two empirical models. The same is true when we consider the role of these
1308 five disturbances for the variance decomposition. Details are given in Tables E.15 and E.16.
1309 One nontrivial change is the role of the productivity disturbance for output growth, which is
1310 16.1% according to SW and 22.2% according to ours. Although the differences seem minor
1311 if we consider the five structural disturbances in isolation, the combined role changes quite a
1312 bit for some variables. For example, the combined role of these five structural disturbances
1313 for investment (amount of capital used) is equal to 55.5% (74.1%) for the SW specification
1314 and 68.7% (92.6%) for our preferred specification.

⁷⁰We report posterior mean estimates unless indicated otherwise.

Table E.14: Posterior Means

| Parameter | Original SW | Agnostic: 2 ASDs | | Agnostic: 3 ASDs | |
|-----------------------|-------------|------------------|--------------|------------------|--------------|
| | | concise | unrestricted | concise | unrestricted |
| α | 0.1903 | 0.2044 | 0.1878 | 0.1877 | 0.2089 |
| σ_c | 1.3889 | 1.4657 | 1.4535 | 1.4618 | 1.4772 |
| Φ | 1.6083 | 1.5211 | 1.5242 | 1.4741 | 1.4762 |
| ϕ | 5.7405 | 5.3843 | 4.4031 | 5.3425 | 4.6933 |
| λ | 0.7136 | 0.6544 | 0.7055 | 0.6679 | 0.6930 |
| ξ_w | 0.7066 | 0.6660 | 0.6706 | 0.7268 | 0.6453 |
| σ_ℓ | 1.8458 | 1.9094 | 1.7733 | 2.0770 | 1.5916 |
| ξ_p | 0.6541 | 0.6566 | 0.6981 | 0.6412 | 0.6902 |
| ι_w | 0.5783 | 0.5556 | 0.5432 | 0.5077 | 0.5557 |
| ι_p | 0.2389 | 0.2010 | 0.1997 | 0.1871 | 0.1891 |
| ψ | 0.5426 | 0.5345 | 0.5049 | 0.5283 | 0.3176 |
| r_π | 2.0469 | 1.7676 | 1.7797 | 1.7746 | 1.7438 |
| ρ | 0.8105 | 0.7933 | 0.8082 | 0.8018 | 0.8032 |
| r_y | 0.0887 | 0.0725 | 0.0860 | 0.0787 | 0.0819 |
| $r_{\Delta y}$ | 0.2237 | 0.1903 | 0.1703 | 0.1941 | 0.1608 |
| ρ_a | 0.9572 | 0.9555 | 0.9483 | 0.9532 | 0.9510 |
| ρ_g | 0.9764 | 0.9719 | 0.9710 | 0.9702 | 0.9018 |
| ρ_r | 0.1464 | 0.1376 | 0.1219 | 0.1286 | 0.1227 |
| ρ_p | 0.8893 | 0.8975 | 0.8899 | 0.9262 | 0.9080 |
| ρ_w | 0.9680 | 0.9751 | 0.9790 | 0.9747 | 0.9822 |
| ρ_b / ρ_A | 0.2165 | 0.3344 | 0.6386 | 0.3239 | 0.4527 |
| ρ_i / ρ_B | 0.7116 | 0.6087 | 0.1660 | 0.6069 | 0.7232 |
| ρ_C | - | - | - | 0.1865 | 0.1577 |
| μ_p | 0.6977 | 0.6764 | 0.6923 | 0.7166 | 0.7172 |
| μ_w | 0.8466 | 0.8241 | 0.8368 | 0.5945 | 0.8168 |
| ρ_{ga} | 0.5184 | 0.6438 | 0.6525 | 0.6709 | 0.5448 |
| σ_a | 0.4586 | 0.4436 | 0.4421 | 0.4524 | 0.4411 |
| σ_g | 0.5299 | 0.4702 | 0.4689 | 0.4428 | 0.2285 |
| σ_r | 0.2449 | 0.2180 | 0.2171 | 0.2171 | 0.2114 |
| σ_p | 0.1403 | 0.1346 | 0.1299 | 0.1308 | 0.1311 |
| σ_w | 0.2427 | 0.2384 | 0.2361 | 0.0763 | 0.2249 |
| σ_b | 0.2398 | - | - | - | - |
| σ_i | 0.4525 | - | - | - | - |
| $100(\beta^{-1} - 1)$ | 0.1648 | 0.1685 | 0.1826 | 0.1656 | 0.2038 |
| $\bar{\gamma}$ | 0.4316 | 0.4349 | 0.4386 | 0.4367 | 0.4352 |
| $\bar{\pi}$ | 0.7845 | 0.7483 | 0.7443 | 0.7391 | 0.7534 |
| $\bar{\ell}$ | 0.5617 | 0.1263 | 0.5216 | 0.1303 | 1.0360 |
| MDD | -922.40 | -892.92 | -906.85 | -890.73 | -925.50 |

Notes. MDD stands for marginal data density. The “concise” ASD specifications are the ones chosen by the specific-to-general model selection procedure. The “unrestricted” ASD specifications are the fully agnostic with no zero restrictions. See Table 1 for the definitions of the parameters.

Table E.15: Variance decomposition for observables across model specifications

| | | ε_a | ε_g | ε_r | ε_p | ε_w | $\varepsilon_b/\tilde{\varepsilon}_A$ | $\varepsilon_i/\tilde{\varepsilon}_B$ | $\tilde{\varepsilon}_C$ |
|------------|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|---------------------------------------|---------------------------------------|-------------------------|
| Δy | Original SW | 16.10 | 28.88 | 6.17 | 4.55 | 6.39 | 22.12 | 15.79 | - |
| | Agnostic: 2 ASDs | 20.29 | 27.01 | 7.15 | 6.04 | 8.12 | 20.53 | 10.85 | - |
| | Agnostic: 3 ASDs | 22.21 | 24.60 | 7.04 | 4.66 | 10.30 | 21.33 | 8.04 | 1.82 |
| Δc | Original SW | 5.29 | 2.10 | 11.56 | 4.40 | 14.54 | 61.17 | 0.95 | - |
| | Agnostic: 2 ASDs | 3.26 | 1.62 | 11.29 | 4.56 | 15.33 | 62.34 | 1.61 | - |
| | Agnostic: 3 ASDs | 2.95 | 1.28 | 10.69 | 3.37 | 17.90 | 61.67 | 2.03 | 0.1 |
| Δi | Original SW | 6.01 | 0.84 | 2.47 | 3.80 | 2.37 | 2.46 | 82.05 | - |
| | Agnostic: 2 ASDs | 4.86 | 0.91 | 2.19 | 4.24 | 2.76 | 12.25 | 72.80 | - |
| | Agnostic: 3 ASDs | 5.49 | 1.02 | 2.38 | 3.80 | 3.94 | 12.55 | 70.01 | 0.81 |
| l | Original SW | 1.94 | 10.34 | 3.15 | 6.23 | 67.66 | 2.52 | 8.15 | - |
| | Agnostic: 2 ASDs | 1.29 | 6.84 | 2.47 | 6.04 | 71.23 | 1.56 | 10.57 | - |
| | Agnostic: 3 ASDs | 1.08 | 4.33 | 2.15 | 4.44 | 79.70 | 1.29 | 4.97 | 2.03 |
| Δw | Original SW | 4.53 | 0.09 | 1.48 | 29.47 | 61.61 | 0.79 | 2.03 | - |
| | Agnostic: 2 ASDs | 3.82 | 0.22 | 2.43 | 30.84 | 54.34 | 3.02 | 5.34 | - |
| | Agnostic: 3 ASDs | 4.09 | 0.11 | 1.25 | 25.18 | 13.32 | 2.23 | 0.38 | 53.45 |
| π | Original SW | 3.92 | 1.00 | 4.25 | 27.64 | 59.43 | 0.58 | 3.18 | - |
| | Agnostic: 2 ASDs | 3.16 | 1.28 | 4.43 | 24.91 | 61.96 | 0.79 | 3.46 | - |
| | Agnostic: 3 ASDs | 2.95 | 0.90 | 3.28 | 16.87 | 70.46 | 0.68 | 3.96 | 0.91 |
| r | Original SW | 10.09 | 3.90 | 14.67 | 7.17 | 38.42 | 7.40 | 18.34 | - |
| | Agnostic: 2 ASDs | 6.50 | 3.49 | 9.77 | 5.79 | 38.96 | 21.49 | 14.02 | - |
| | Agnostic: 3 ASDs | 5.70 | 2.77 | 8.18 | 4.33 | 48.61 | 17.29 | 12.47 | 0.65 |

Notes. The table provides the contributions (in percent) of the different structural disturbances to the variance of the observable variables, across different model specifications. The ASD specifications are the ones chosen by our model selection procedure. y stands for log output; c for log consumption; i for log investment; l for hours; w for log wage rate; π for inflation; and r for nominal interest rate. Structural disturbances are defined as follows. ε_a : TFP; ε_g : government expenditures; ε_r : monetary policy; ε_p : price mark-up; ε_w : wage mark-up; ε_b : risk premium; ε_i : investment; $\tilde{\varepsilon}_A$: agnostic Euler; $\tilde{\varepsilon}_B$: agnostic investment-modernization; and $\tilde{\varepsilon}_C$: capital-efficiency wage mark-up.

Table E.16: Variance decomposition for additional variables across model specifications

| | | ε_a | ε_g | ε_r | ε_p | ε_w | $\varepsilon_b/\tilde{\varepsilon}_A$ | $\varepsilon_i/\tilde{\varepsilon}_B$ | $\tilde{\varepsilon}_C$ |
|-----------|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|---------------------------------------|---------------------------------------|-------------------------|
| y_t | Original SW | 29.93 | 4.09 | 2.16 | 6.37 | 48.58 | 1.53 | 7.34 | - |
| | Agnostic: 2 ASDs | 26.50 | 3.02 | 1.91 | 7.02 | 55.93 | 1.31 | 4.32 | - |
| | Agnostic: 3 ASDs | 21.19 | 2.13 | 1.67 | 5.47 | 65.95 | 1.14 | 2.17 | 0.28 |
| c_t | Original SW | 11.06 | 8.42 | 2.08 | 4.19 | 69.25 | 2.18 | 2.83 | - |
| | Agnostic: 2 ASDs | 6.60 | 6.60 | 1.78 | 4.23 | 78.76 | 1.81 | 0.22 | - |
| | Agnostic: 3 ASDs | 4.29 | 4.30 | 1.52 | 3.16 | 84.48 | 1.51 | 0.49 | 0.25 |
| i_t | Original SW | 20.37 | 5.41 | 1.27 | 6.93 | 21.56 | 0.22 | 44.23 | - |
| | Agnostic: 2 ASDs | 17.22 | 6.35 | 1.14 | 8.25 | 29.78 | 1.21 | 36.04 | - |
| | Agnostic: 3 ASDs | 15.31 | 5.79 | 1.13 | 7.75 | 38.72 | 1.06 | 29.25 | 1.00 |
| r_t^k | Original SW | 14.86 | 17.47 | 1.63 | 10.58 | 19.21 | 0.86 | 35.39 | - |
| | Agnostic: 2 ASDs | 12.28 | 20.44 | 2.65 | 19.09 | 29.73 | 0.92 | 14.88 | - |
| | Agnostic: 3 ASDs | 8.41 | 14.66 | 1.73 | 13.16 | 30.17 | 0.67 | 18.12 | 13.08 |
| q_t | Original SW | 4.65 | 0.55 | 9.03 | 3.11 | 1.20 | 45.42 | 36.04 | - |
| | Agnostic: 2 ASDs | 9.78 | 1.35 | 21.83 | 9.30 | 3.67 | 19.58 | 34.49 | - |
| | Agnostic: 3 ASDs | 9.72 | 1.35 | 19.88 | 6.50 | 5.14 | 18.64 | 31.56 | 7.21 |
| z_t | Original SW | 14.86 | 17.47 | 1.63 | 10.58 | 19.21 | 0.86 | 35.39 | - |
| | Agnostic: 2 ASDs | 12.23 | 20.36 | 2.64 | 19.02 | 29.61 | 4.43 | 11.71 | - |
| | Agnostic: 3 ASDs | 8.86 | 15.43 | 1.82 | 13.85 | 31.76 | 4.14 | 9.46 | 14.68 |
| μ_t^p | Original SW | 11.56 | 0.29 | 3.27 | 57.02 | 23.87 | 0.87 | 3.11 | - |
| | Agnostic: 2 ASDs | 8.06 | 0.37 | 3.38 | 53.22 | 18.90 | 14.59 | 1.48 | - |
| | Agnostic: 3 ASDs | 7.99 | 0.24 | 2.13 | 54.80 | 11.88 | 15.22 | 2.61 | 5.13 |
| k_t^s | Original SW | 23.43 | 3.92 | 1.23 | 11.37 | 34.19 | 0.36 | 25.50 | - |
| | Agnostic: 2 ASDs | 21.82 | 4.90 | 1.55 | 16.51 | 52.66 | 1.32 | 1.24 | - |
| | Agnostic: 3 ASDs | 15.59 | 3.60 | 1.11 | 14.11 | 58.20 | 1.21 | 0.61 | 5.57 |
| k_t | Original SW | 22.38 | 8.11 | 0.50 | 4.93 | 31.56 | 0.04 | 32.48 | - |
| | Agnostic: 2 ASDs | 22.16 | 11.30 | 0.55 | 7.27 | 55.74 | 0.21 | 2.77 | - |
| | Agnostic: 3 ASDs | 14.84 | 8.05 | 0.42 | 6.18 | 58.26 | 0.12 | 2.37 | 9.75 |
| w_t | Original SW | 33.03 | 1.03 | 1.95 | 38.38 | 18.61 | 0.40 | 6.60 | - |
| | Agnostic: 2 ASDs | 26.99 | 1.00 | 2.63 | 47.34 | 20.71 | 0.39 | 0.92 | - |
| | Agnostic: 3 ASDs | 25.35 | 0.74 | 1.62 | 49.34 | 14.29 | 0.30 | 0.44 | 7.92 |

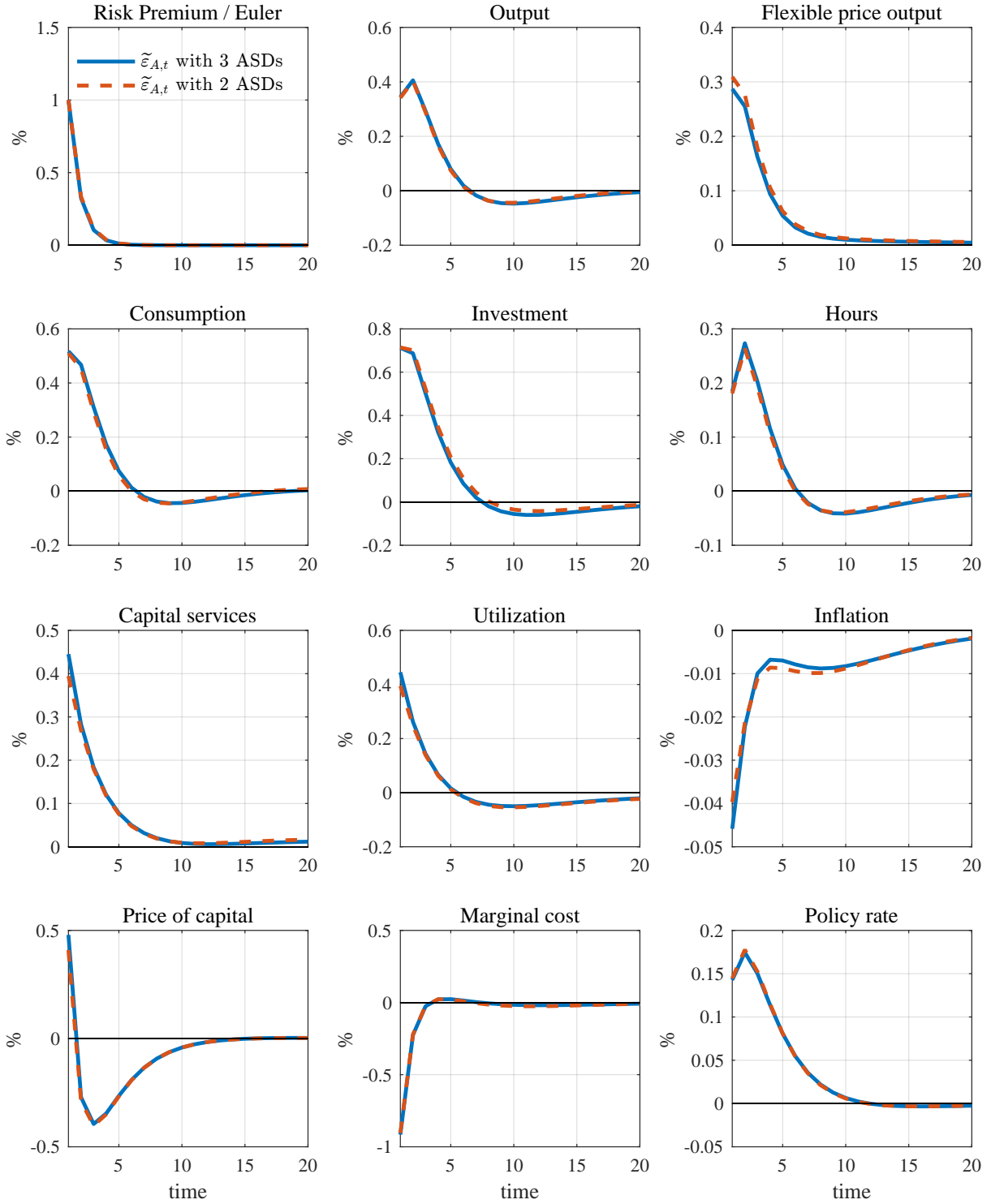
Notes. The table provides the contributions (in percent) of the different structural disturbances to the variance of the observable variables, across different model specifications. The ASD specifications are the ones chosen by our model selection procedure. y stands for log output; c for log consumption; i for log investment; l for hours; w for log wage rate; r^k for rental rate on capital; q for the log price of capital; z for the utilization rate; μ^p for the price mark-up; k^s for log capital used in production; and k for log installed capital. Structural disturbances are defined as follows. ε_a : TFP; ε_g : government expenditures; ε_r : monetary policy; ε_p : price mark-up; ε_w : wage mark-up; ε_b : risk premium; ε_i : investment; $\tilde{\varepsilon}_A$: agnostic Euler; $\tilde{\varepsilon}_B$: agnostic investment-modernization; and $\tilde{\varepsilon}_C$: capital-efficiency wage mark-up.

1315 **Specifications with and without restrictions on ASDs.** Table E.14 also compares
1316 structural parameter estimates of concise ASD models chosen by our model selection proce-
1317 dures with those that still allow ASDs to enter all equations. The parameter estimates are
1318 fairly similar. IRFs for the included regular structural disturbances are also quite similar.

1319 That is not always the case for the IRFs of the agnostic disturbances themselves. The IRFs
1320 for some variables do differ between the concise and the fully unrestricted ASD specification.
1321 Given the misspecification results of Appendix D, it is not surprising that different empirical
1322 specifications lead to different results. Another issue with the fully unrestricted ASD spec-
1323 ification is that it estimates a large number of coefficients which complicates generating an
1324 accurate posterior with Monte Carlo Markov Chain algorithms. Especially, for the 3-ASD
1325 fully unrestricted specification, the Brooks-Gelman statistics did not look particularly good
1326 for some of the coefficients associated with the agnostic disturbances. Thus, we prefer the
1327 concise ASD specifications.

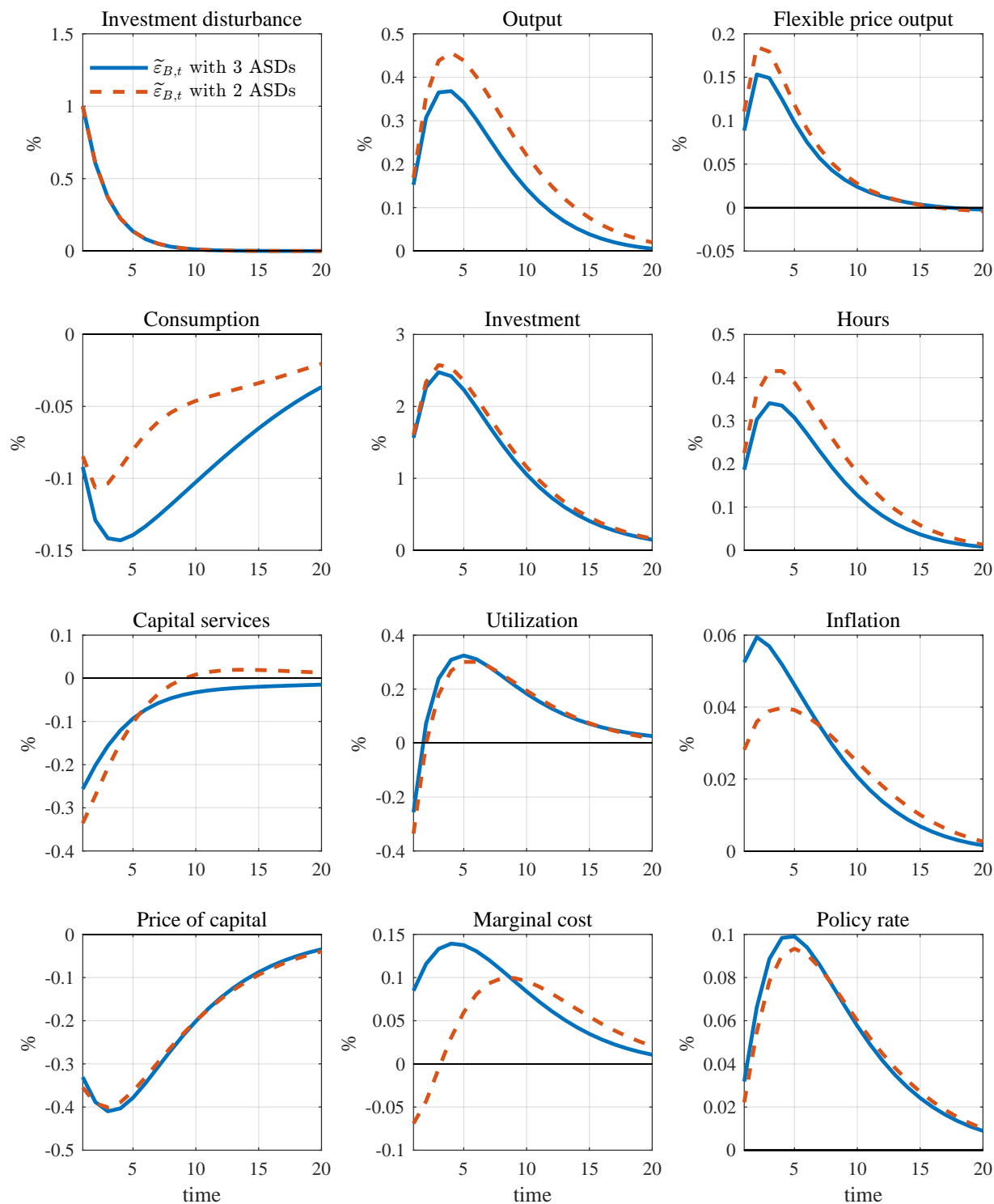
1328 **Specifications with two and three ASDs.** Tables E.15 and E.16 provide the role of
1329 the regular and agnostic disturbances for the fluctuations of a wide range of variables. In
1330 addition to the results of the SW specification, it also shows the results for the two-ASD and
1331 three-ASD specification chosen by our specific-to-general model selection procedure. It shows
1332 that the results are very similar for the two chosen ASD specifications. The same conclusion
1333 can be drawn from Figures E.9 and E.10 that plot the IRFs for two agnostic disturbances.

Figure E.9: IRFs of the agnostic Euler disturbance: 2 versus 3 ASDs



Notes. These panels plot the IRFs of the agnostic disturbance $\tilde{\varepsilon}_{A,t}$ that we interpret as a general Euler disturbance for the empirical specifications with two and three ASDs. Both specifications are chosen with the specific-to-general model selection procedure.

Figure E.10: IRFs of the agnostic investment-modernization disturbance: 2 versus 3 ASDs



Notes. These panels plot the IRFs of the agnostic disturbance $\tilde{\varepsilon}_{B,t}$ that we interpret as an investment-modernization disturbance for the empirical specifications with two and three ASDs. Both specifications are chosen with the specific-to-general model selection procedure.

1334 **Additional results for $\tilde{\varepsilon}_{A,t}$.** Figure E.11 plots the IRFs associated with an innovation
1335 in the agnostic Euler disturbance for our 3-ASD benchmark specification and also when the
1336 coefficient of this agnostic disturbance in the capital valuation equation is equal to zero. A
1337 preference disturbance does not show up in this equation and a bond risk-premium distur-
1338 bance does.⁷¹ The IRFs are very similar, which confirms our claim that the coefficient in the
1339 capital valuation equation is quantitatively not very important. This does not mean that
1340 the ASD is a preference disturbance, since the ASD shows up in the investment equation
1341 whereas a preference disturbance does not.

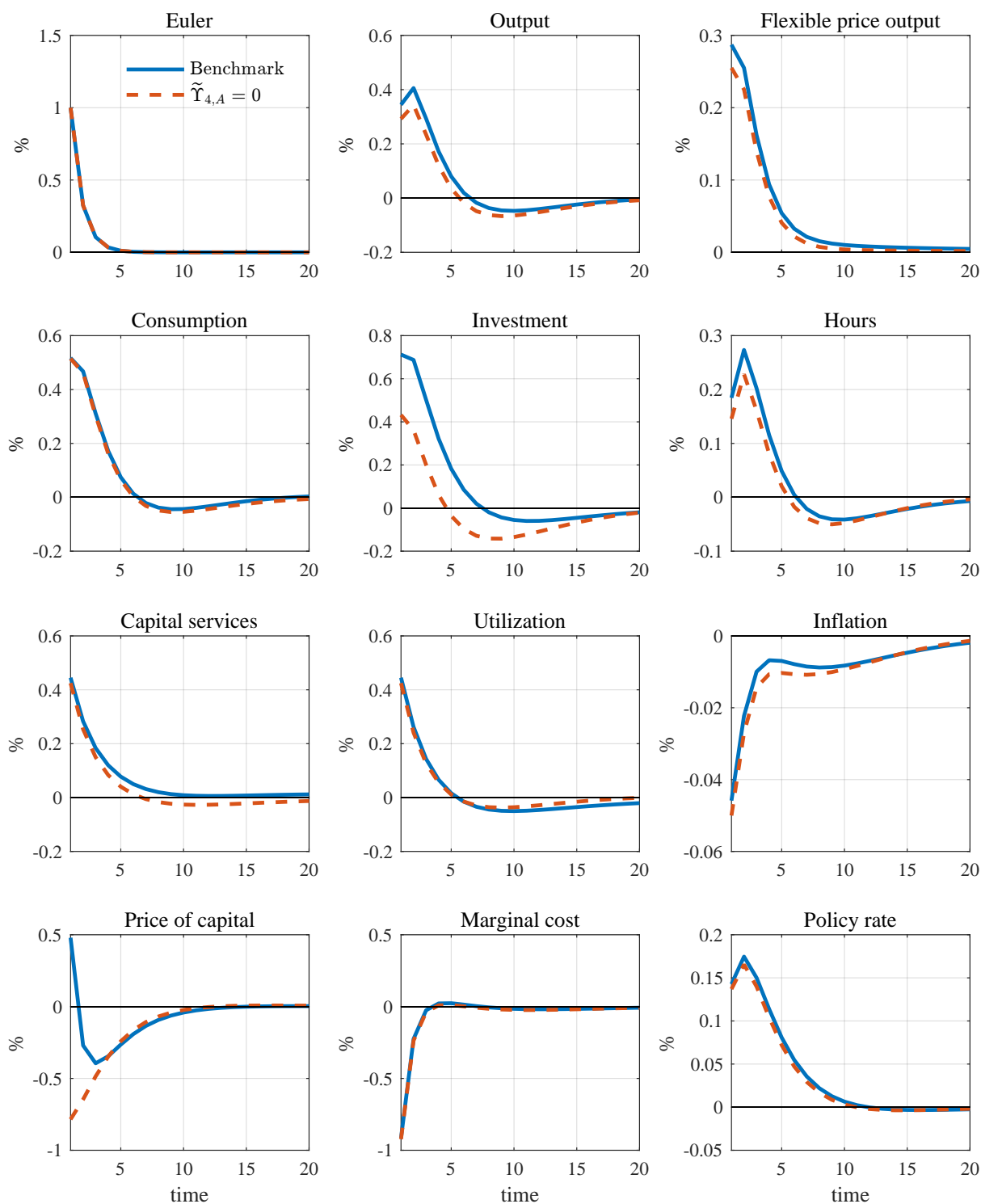
1342 Figure E.12 plots the same IRFs when the coefficient of the agnostic Euler disturbance in
1343 the Taylor rule is set equal to zero. The figure shows that the direct response of the policy
1344 rate to a positive shock to this disturbance dampens the expansion and prevents an upsurge
1345 of inflation.

1346 Figure E.13 plots the same IRFs when we set equal to zero the coefficients of the dis-
1347 turbance in the four equations that we ignored in the discussion of the agnostic Euler dis-
1348 turbance, namely, the overall budget constraint, the utilization, the price mark-up equation,
1349 and the rental rate of capital equation. The figure documents that the role of the agnostic
1350 disturbance through these equations is minor since the IRFs are overall quite similar to those
1351 of our benchmark specification.

1352 **Additional results for $\tilde{\varepsilon}_{C,t}$.** Figure E.14 plots the IRFs for our agnostic capital-efficiency
1353 wage mark-up disturbance when the coefficient of this disturbance in the overall budget
1354 constraint is set equal to zero. The figure documents that this has a minor impact on IRFs.

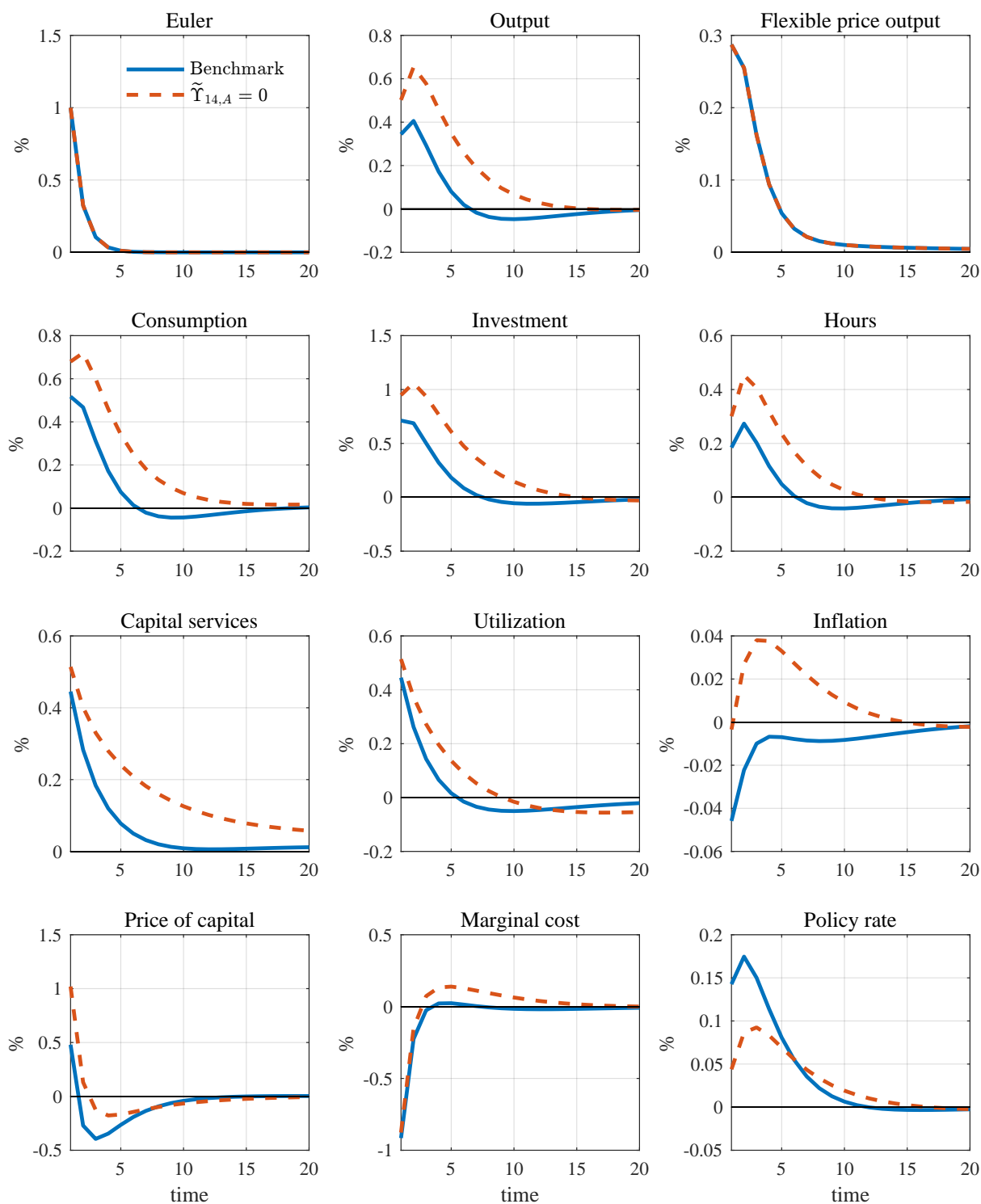
⁷¹Recall that the MRS has been substituted out of the capital valuation equation using the MRS of the bond Euler equation.

Figure E.11: IRFs of the agnostic Euler disturbance with restrictions I



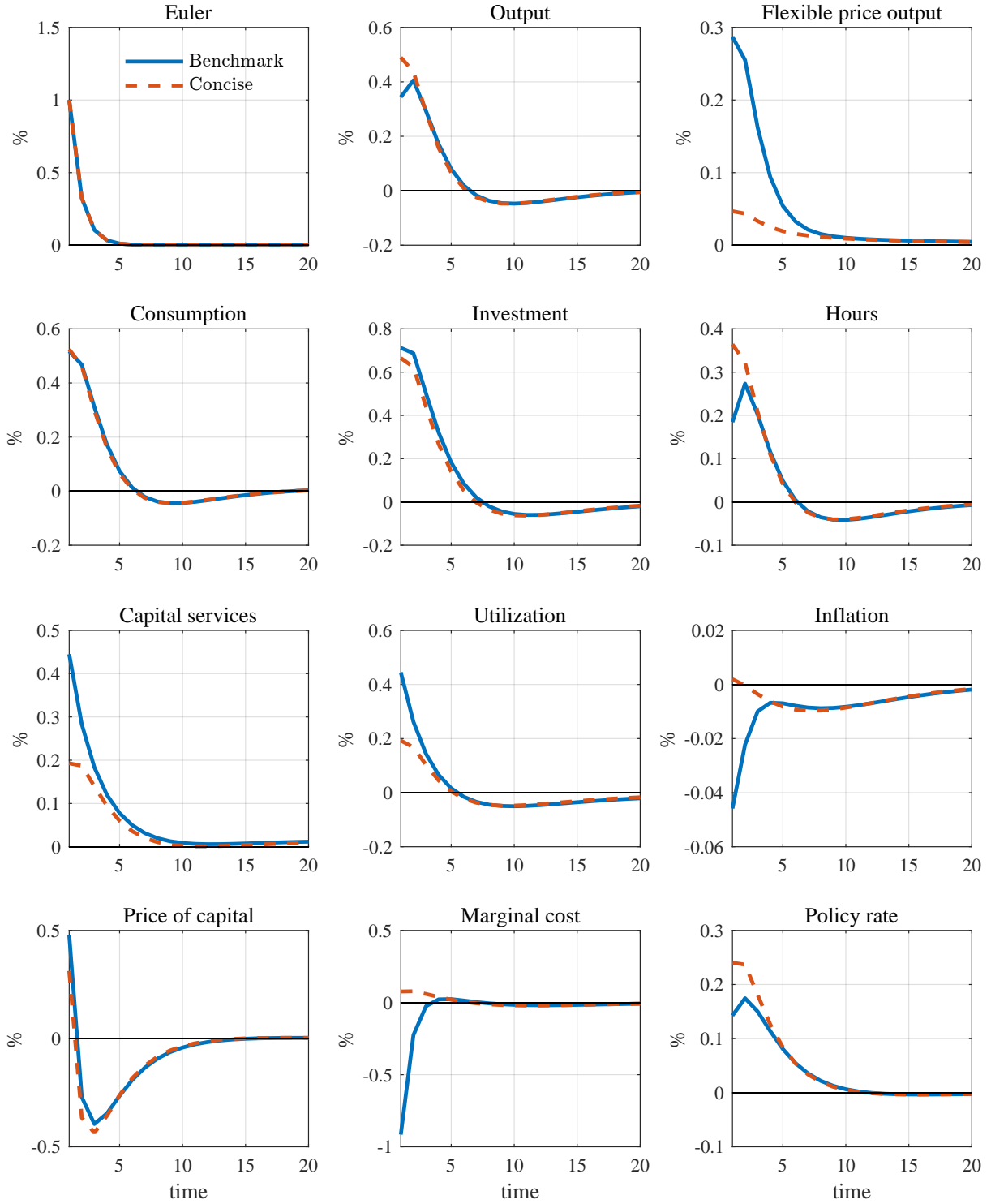
Notes. These panels plot the IRFs of the agnostic Euler disturbance for our benchmark specification and when the impact of this IRF through the capital valuation equation is set equal to zero.

Figure E.12: IRFs of the agnostic Euler disturbance with restrictions II



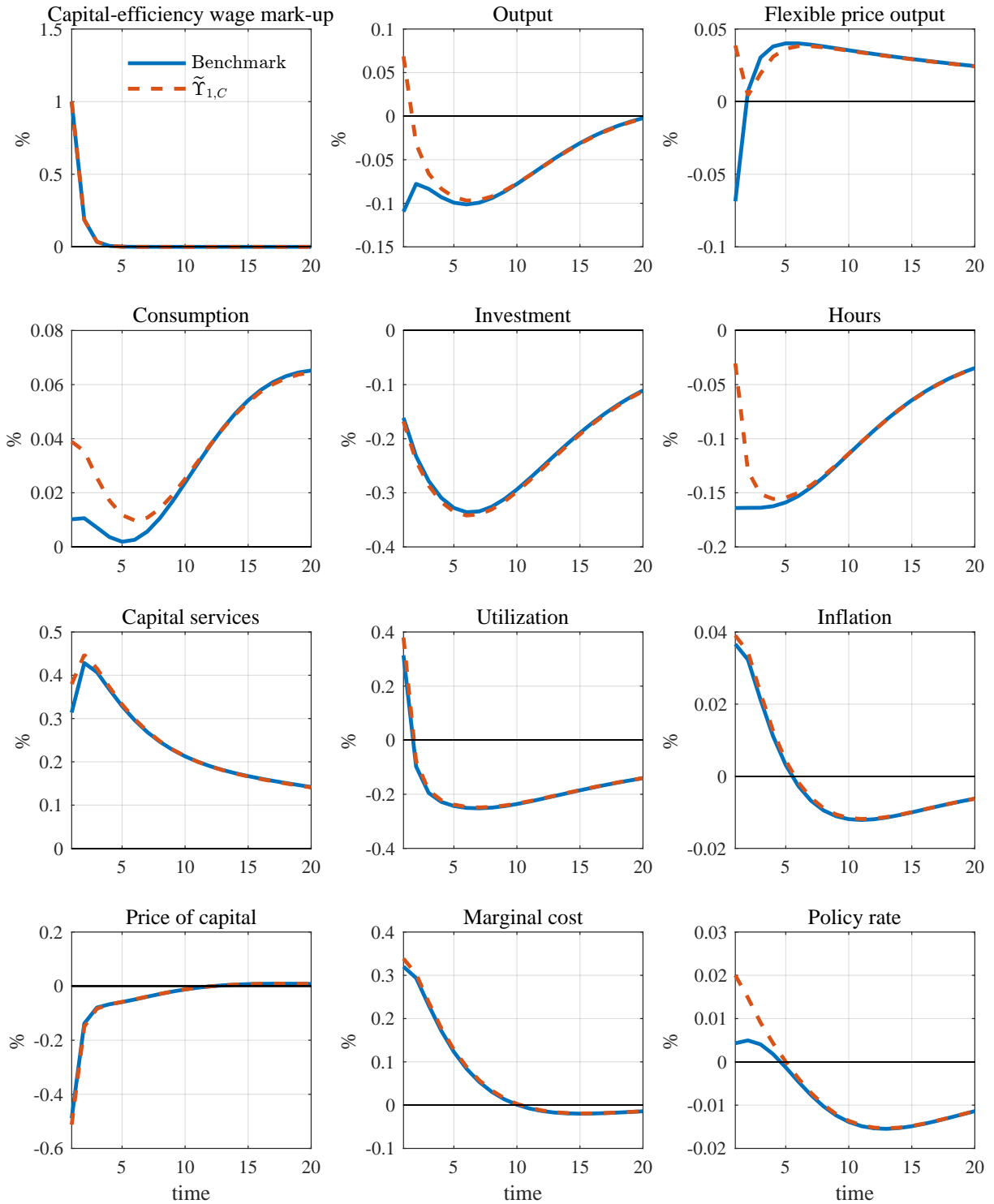
Notes. These panels plot the IRFs of the agnostic Euler disturbance for our benchmark specification and when the impact of this IRF through the Taylor rule is set equal to zero.

Figure E.13: IRFs of the agnostic Euler disturbance with restrictions III



Notes. These panels plot the IRFs of the agnostic Euler disturbance for our benchmark specification and when the impact of this IRF through the overall budget constraint, the utilization, the price mark-up equation, and the rental rate of capital equation is set equal to zero.

Figure E.14: IRFs of the agnostic capital-efficiency wage mark-up disturbance with restrictions



Notes. These panels plot the IRFs of the agnostic capital-efficiency wage mark-up disturbance for our benchmark specification and when the impact of this IRF through the overall budget constraint is set equal to zero.