

Identifying Technology Shocks in Models with Heterogeneous Inputs

Luigi Bocola*
University of Pennsylvania

Marcus Hagedorn†
University of Zurich

Iourii Manovskii‡
University of Pennsylvania

March 31, 2011

Preliminary and Incomplete

Abstract

We show that the key identifying assumptions underlying the existing approaches to estimating neutral aggregate technology shocks are violated in models with heterogeneous capital and labor. We propose a new method to identifying technology shocks in the data in the presence of factor heterogeneity and prove its identification. The results of Monte Carlo simulations imply that the proposed method has better small sample properties than the existing methods. We apply the proposed method to assess the role of neutral technology in driving growth and fluctuations in US data.

JEL Classification: E32, E24

Keywords: Economic Shocks, Heterogeneity, Growth Accounting, Business Cycles

*Department of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA, 19104-6297 USA. E-mail: lbocola@sas.upenn.edu.

†Institute for Empirical Research (IEW), University of Zurich, Blümlisalpstrasse 10, CH-8006 Zürich, Switzerland. Email: hagedorn@iew.uzh.ch.

‡Department of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA, 19104-6297 USA. E-mail: manovski@econ.upenn.edu.

1 Introduction

Neutral technical change is of central importance to researchers of economic growth and business cycle fluctuations. For understanding the sources of economic growth over particular historical episodes it is often essential to separate the growth in efficiency of particular inputs used in production from the Hicks-neutral, or disembodied, technological change that does not affect the marginal rate of substitution between factors of production. The evolution of technology at business cycle frequencies is essential for understanding the sources of economic fluctuations. Moreover, the relationships between neutral technology and other economic variables provide the essential guidance for the development of economic models. These relationships are also routinely used to distinguish between competing models. For example, the empirical finding that aggregate hours worked fall in response to a technology shock called into question the usefulness of the Real-business-cycle (RBC) model (Kydland and Prescott (1982), Long and Plosser (1983)) for interpreting aggregate fluctuations.

In light of the importance of the technology series, considerable effort was devoted to obtaining robust empirical measures of the (growth of) neutral technology. Yet, in this paper we will argue that the key identifying assumptions of the existing empirical approaches are typically not satisfied in the presence of heterogeneous capital and labor inputs (e.g., when the effective labor input aggregates services of different and possibly imperfectly substitutable labor inputs, e.g., high and low skilled, young and old, workers in white collar or blue collar occupations, etc). This leads us to develop a new method to estimate neutral technology that is robust to the presence of (unobserved) factor heterogeneity.

The classic approach proposed by Solow (1957) identified neutral technology with the residual output growth, i.e., the growth of output not accounted for by the growth of inputs.¹ Clearly, because the technology shock is computed as a residual, the contribution of all the factors of production must be accounted for which places high demands on the data. Moreover, to identify technology shocks following this approach it is essential that an econometrician knows the true production function. If the production function used by the econometrician does not aggregate heterogeneous inputs appropriately, the identified technology shocks will

¹The current state-of-the-art implementation of this procedure can be found in Basu, Fernald, and Kimball (2006).

be biased. Unfortunately, the true aggregation of heterogeneous labor and capital inputs is not known. The solution adopted in the literature, following the work by Jorgenson, is to aggregate all distinct labor and capital inputs into aggregate labor and capital inputs measured in efficiency units using Tornqvist indices.² The bias of this procedure is small if (1) the distinct factor inputs observed in the data correspond to true distinct inputs into the aggregate production function, (2) the true production function is well approximated by a translog function, and (3) the parameters of the aggregate production function (for example, the relative productivity of different inputs) are constant over time. The latter requirement represents a major empirical challenge to this framework as it is well documented that relative factor productivities exhibit non-trivial dynamics (e.g., Katz and Murphy (1992) interpret the skill-biased technical change as the increase in the relative productivity of highly educated and experienced workers). We will show that a violation of this assumption imparts a quantitatively sizable bias on technology shocks identified as Solow residuals with Jorgenson’s correction.

An alternative approach to identifying neutral technology shocks that drew substantial recent interest in the literature is based on the identifying assumption that only technology shocks have a long-run effect on labor productivity.³ The idea of using long-run restrictions was developed in Blanchard and Quah (1989) and King, Plosser, Stock, and Watson (1991) and was implemented in the business cycle context by Gali (1999) using structural vector autoregressions (SVAR). Although this approach has been intensely discussed in the literature, the key identifying assumption is rarely questioned.⁴ Instead, most of the issues raised are of

²This is also the approach followed in practice. For example, the measures of productivity provided by the Bureau of Labor Statistics are based on this procedure.

³A third approach adopts a more narrow interpretation of productivity shocks and attempts to directly measure technological advances in the data. In a business cycle context, for example, Shea (1998) finds that technological innovations identified using observations on R&D and patents in the data have only a weak relationship to TFP and hours. In contrast, Alexopoulos (2010) finds that technological innovations identified using data on books published in the field of technology are strongly positively correlated with TFP and hours of work.

⁴It is widely recognized that this assumption is inconsistent with endogenous growth models. Capital taxes also may have a long-run effect on labor productivity in a RBC model but were found to have a relatively small effect in Francis and Ramey (2005). Shea (1998) suggests that if low-productivity firms are destroyed in recessions this might have a long run effect on productivity. Uhlig (2004) argues that permanently changing

econometric nature, such as the small sample bias, small number of lags, etc (e.g., Faust and Leeper (1997), Chari, Kehoe, and McGrattan (2008), Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007)). In this paper we question the long-run restriction itself. In particular, we show that non-technology shocks do not have a long-run effect on productivity only in a very restrictive class of models with homogeneous capital and homogeneous labor. In models with heterogeneous labor or capital inputs most non-technology shocks can have a long-run effect on labor productivity as well. Thus, even if econometric difficulties associated with this approach can be overcome,⁵ the fundamental economic difficulties presented by factor heterogeneity will lead to biased inference of neutral technology.

Given these limitations of the existing procedures we propose an alternative method for estimating neutral technology shocks that is robust to the presence of factor heterogeneity. The proposed method does not make use of the long run restrictions and does not impute the technology shocks as a residual. Instead, we interpret the technology shock as an unobserved state and, using some equilibrium conditions of the model, we prove that this state can be identified using filtering/smoothing techniques. Since we do not treat the technology shock as a residual our method does not require to specify an explicit function that aggregates heterogeneous labor inputs. Moreover, it does not require the parameters of this function to be invariant over time. Rather, the derivatives of that function are themselves treated as unobserved states in the filtering problem.

While heterogeneity presents a challenge to the existing procedures, it enables the identification of shocks in the method proposed in this paper. For example, the log of the wage of workers of a particular type can be written as a sum of a common component and a component that is partially idiosyncratic to that worker type. The system of equations represented by wages of at least two distinct labor inputs thus has a factor structure that we exploit to identify the common neutral technology component.

The paper is organized as follows. In Section 2 we describe the procedures used in the social attitudes to workplace whereby workers substitute leisure activities at home with leisure activities at work will result in effective mismeasurement of effective work hours and affect measured productivity in the long run. The latter two papers do not study the quantitative importance of these mechanisms.

⁵For example, Christiano, Eichenbaum, and Vigfusson (2006) propose alternative procedures for estimating long-run-identified SVARs with better small sample properties.

literature to identify technology shocks and show theoretically that they are biased in the presence of worker heterogeneity that is not properly accounted for. In Section 3 we develop a new method to recover technology shocks and prove its identification. In Section 4 we assess how the different methodologies perform in small samples drawn from an estimated RBC model with heterogeneous labor. The results suggest that the method that we propose performs substantially better than the existing ones even in small samples. Finally, in Section 5 we apply our method in the data and estimate a quarterly technology series for the US. We describe and analyze the sequence of identified shocks and document its co-movement with other economic aggregates. Section 6 concludes.

2 Inputs' Heterogeneity and the Identification of Technology Shocks

2.1 Identification of technology shocks in models with homogeneous inputs

The environment we consider is that of an RBC model with homogeneous inputs. We assume that households are all alike and value consumption, c_t , and dislike labor, h_t , according to a following utility function:

$$U(c_t, h_t, A_t). \tag{1}$$

A_t represents a shock to preferences and it is best thought as a reduced form representation of the “demand” disturbances that affect individual’s attitude toward working in the market or not. It summarizes the effects of productivity shocks in the home production technology relative to market technology as in Benhabib, Rogerson, and Wright (1991), changes in the share in government spending as in Christiano and Eichenbaum (1988), true preference shocks as in Bencivenga (1992), changes in labor income tax and transfers policies as in Prescott (2004), etc. We assume that a final homogeneous good Y_t can be produced in the economy according to the CRS technology

$$Y_t = F(K_t, e^{Z_t} L_t), \tag{2}$$

where K_t and L_t denote the effective capital and labor-input services employed (even allowing for unobservable variations in the utilization rate of both inputs) and Z_t is an exogenous stochastic technology parameter.

A typical household enters period t with some savings, k_t , remunerated at the gross rate of R_t , it earns a wage W_t for the hours worked h_t , and it spends its total income to acquire the final good Y_t which is then used for consumption and saving purposes. It is assumed that all these choices are optimal, in the sense that they solve the decision problem:

$$\begin{aligned} & \max_{\{c_t, h_t, k_{t+1}\}_{t=0}^{\infty}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t U(c_t, h_t, A_t) \right\} \\ & c_t + k_{t+1} \leq W_t h_t + R_t k_t \quad \forall t \\ & c_t \geq 0 \quad k_{t+1} \geq 0 \quad h_t \in [0, 1] \quad \forall t \end{aligned}$$

where the stochastic process governing A_t and Z_t , wages W_t and capital rental rates R_t are taken as given by the household. A typical firm at time t rents labor and capital services from the households sector at competitive prices and use the technology of Equation (2) to produce the final output Y_t . Factors' demand at time t is determined as the solution to the static optimization program:

$$\max_{K_t \geq 0, L_t \geq 0} \left[F(K_t, e^{Z_t} L_t) - W_t L_t - R_t K_t \right]. \quad (3)$$

The model is completed by a stochastic law of motion for (Z_t, A_t) and by the assumption that factors' prices clear the markets for capital and labor services in every period.

This set of conditions concerning preferences, technology, agents behavior and the functioning of markets implicitly defines the stochastic law of motion for the endogenous variables: real GDP, hours worked, real wages, etc. at time T are all functions of the path for the exogenous disturbances $\{A_t, Z_t\}_{t=1}^T$ and, as such, they carry information on the unobservable structural shocks. Based on this insight, researchers have thought of using some of the restrictions implied by RBC models to back-out the actual realization of technology from aggregate time series. Different procedures have been developed in the literature for this purpose.

A first approach consists in calculating the technology series as the residual of an aggregate production function. We can use a log-linear approximation of the technology in Equation

(2) and first difference the variables to obtain:

$$\Delta \log(Y_t) = \alpha \Delta \log(K_t) + (1 - \alpha) \Delta \log(L_t) + (1 - \alpha) \Delta Z_t. \quad (4)$$

This approximation would be exact if α and $1 - \alpha$ are the constant factor shares and thus the technology is Cobb-Douglas. Since output, hours worked and capital are observed in the data, knowledge of α would allow the researcher to recover the realization of the technology series $\{\Delta Z_t\}$. While conceptually straightforward, this procedure poses major econometric challenges. For example, the RBC model suggests that the labor input reacts to contemporaneous changes in technology, making the regressors in Equation (4) endogenous.⁶ Moreover, varying utilization rates of capital and labor, if not properly accounted, would be reflected in the Solow residual and would lead to a mismeasurement of the resulting technology series. Basu, Fernald, and Kimball (2006) address these and related issues and back-out a quarterly technology series for the US finding that hours worked respond negatively to a technology shock.

In view of the difficulties that arise in the estimation of aggregate production functions, researchers have thought of using other approaches to identify technology shocks. In particular, Gali (1999) points out that a wide range of RBC models share a common long-run neutrality property: *only* technology shocks have a long-run effect on labor productivity. In our model, for example, the logarithm of output per worker is given by:

$$\log \left(\frac{Y_t}{L_t} \right) = Z_t + \log \left(F \left(\frac{K_t}{e^{Z_t} L_t}, 1 \right) \right). \quad (5)$$

Under the auxiliary assumption that capital intensity $\frac{K_t}{e^{Z_t} L_t}$ follows a stationary stochastic process (or equivalently the interest rate in the economy is a stationary series), it is not affected in the long run by any structural shock. Thus, Equation (5) suggests that stochastic trends in labor productivity can be induced *only* by the technology shock Z_t . The idea of Gali (1999), then, is to retrieve the technology series from the low frequency behavior of $\log \left(\frac{Y_t}{L_t} \right)$.

⁶In principle, one could get consistent estimates for capital and labor elasticities using information on labor income share, thus avoiding this endogeneity problem. Though, this is true only for the Cobb-Douglas case and only if we were to impose that factor prices equal marginal productivities at every point in time. Since the objective of this literature is to identify technology shocks placing as little theoretical restrictions as possible, researchers typically end up estimating the production function using instrumental variables.

In particular, applying the procedure developed in Blanchard and Quah (1989) on a bivariate VAR for labor productivity and hours worked, Gali (1999) distinguishes between shocks that have a “long run” effect on $\log\left(\frac{Y_t}{L_t}\right)$ and shocks that have only a transitory effects on it: the former, then, are interpreted as a direct measure for technology. In his empirical application, Gali (1999) finds that hours worked react negatively to the identified technology shock.

The findings of Basu, Fernald, and Kimball (2006) and Gali (1999) have been seen as a major counterfactual of the RBC model and they have been hotly debated in the literature. Most of the focus has been placed on the empirical strategy adopted by Gali (1999). In particular, researchers have mostly questioned the use of SVARs for implementing the long run identification scheme. The growth accounting procedure, however, does not suffer from the econometric issues highlighted above and, therefore, it can potentially validate the robustness of the SVAR approach. In fact, Gali (2004) has recently showed that the technology series identified using the SVAR and the production function procedure implemented in Basu, Fernald, and Kimball (2006) are highly correlated. This correlation is interpreted as a sign of robustness of Gali (1999) findings to the criticisms raised in the literature.

In what follows we show that the identifying restrictions underlying the production function approach and the SVAR procedure with long run restrictions are not valid when the data generating process features a particular form of workers’ heterogeneity. We also show that under a reasonable restriction of the structural parameters, both procedures generate a systematic downward bias in the estimation of the response of hours to technology shocks.

2.2 The Effects of Input Heterogeneity

We introduce workers’ heterogeneity in its simplest form, assuming that there are two types of households in our economy, indexed by u and s (e.g., unskilled and skilled). The total measure of households in the economy is $1 = u + s$, i.e., the sum of measures of unskilled and skilled households. Suppose agents of type $j = \{u, s\}$ value consumption and leisure according to type-dependent utility functions:

$$U_j(c_t, h_t, A_t). \tag{6}$$

We are also going to assume that firms use the production function:

$$Y_t = F(K_t, e^{Z_t} L_{s,t}, e^{Z_t} L_{u,t}), \quad (7)$$

where $L_{s,t} = sh_{s,t}$ is total hours worked by skilled individuals, $L_{u,t} = uh_{u,t}$ is total hours worked by unskilled individuals, and $L_t = L_{s,t} + L_{u,t}$ is total hours worked. Beside these modifications, the model is the one described in the previous section.

The main question we ask in this section is what happens in this environment if a researcher backs-out technology shocks as the residual of a production function or as the shock affecting labor productivity in the long run. Notice, first, that the simple Solow residual in this economy no longer represents a direct measure of technology. From the production function in equation (7), in fact, we can use a log-linear approximation of the technology (7) and first difference the variables to obtain:

$$\Delta \log(Y_t) = \alpha \Delta \log(K_t) + \alpha_u \Delta \log(L_{u,t}) + \alpha_s \Delta \log(L_{s,t}) + (1 - \alpha) \Delta Z_t, \quad (8)$$

where CRS imply that $\alpha_u + \alpha_s = 1 - \alpha$. We get that

$$\begin{aligned} & \Delta \log(Y_t) - \alpha \Delta \log(K_t) - (1 - \alpha) \Delta \log(L_t) \\ &= (1 - \alpha) \Delta Z_t + \alpha_u \Delta \log\left(\frac{L_{u,t}}{L_t}\right) + \alpha_s \Delta \log\left(\frac{L_{s,t}}{L_t}\right) \\ &= (1 - \alpha) \left[\Delta Z_t + \Delta \log\left(\frac{L_{e,t}}{L_t}\right) \right], \end{aligned} \quad (9)$$

where L_t^e , effective labor input, is an aggregator of unskilled and skilled labor inputs $L_t^{e,t} = L_{u,t}^{\frac{\alpha_u}{1-\alpha}} L_{s,t}^{\frac{\alpha_s}{1-\alpha}}$.

The failure to properly measure hours worked in efficiency units, thus, results in an improper account of technology if shocks, for example to the disutility of labor, have an impact on $\log\left(\frac{L_t^e}{L_t}\right)$ and are consequently reflected in the Solow residual.

We can also show that the long-run neutrality property of labor productivity is lost in this model. In fact, the logarithm of labor productivity in this economy is given by:

$$\log\left(\frac{Y_t}{L_t}\right) = Z_t + \log F\left(\frac{K_t}{e^{Z_t} L_t}, \frac{L_{s,t}}{L_t}, \frac{L_{u,t}}{L_t}\right).$$

If neutral technology shocks were the only shocks affecting labor productivity in the long run

then

$$\underbrace{\frac{Y_t}{e^{Z_t} L_t}}_{stationary} = F\left(\underbrace{\frac{K_t}{e^{Z_t} L_t}}_{stationary}, \frac{L_{s,t}}{L_t}, \frac{L_{u,t}}{L_t}\right).$$

This implies that any non-neutral technology shock X does not affect the LHS in the long-run and thus does not affect the RHS. Since $\frac{K_t}{e^{Z_t} L_t}$ is also stationary it is not affected by X in the long-run either. As a result, the long-run changes of L_u and L_s in response to a shock to X have to neutralize each other:

$$F_{L_s} \frac{\partial \frac{L_s}{L}}{\partial X} + F_{L_u} \frac{\partial \frac{L_u}{L}}{\partial X} = \frac{\partial \frac{L_s}{L}}{\partial X} (F_{L_s} - F_{L_u}) = 0. \quad (10)$$

We have therefore established the following proposition:

Proposition 1. *Let $L_s(w_s, A)$ and $L_u(w_u, A)$ be labor supply of the two groups. If for the elasticities $\epsilon_{L_s, A} \neq \epsilon_{L_u, A}$ and if the two groups have distinct marginal products $F_{L_s} \neq F_{L_u}$ then $\frac{Y_t}{e^{Z_t} L_t}$ is not stationary. In particular, any shock that changes labor composition in the long-run, has a permanent effect on labor productivity $\frac{Y_t}{L_t}$.*

The implication of Proposition 1 is that any long-run restriction on labor productivity will spuriously interpret persistent movements in demand disturbances as a technology shock. In order to gain intuition for this result consider a shock to the disutility of labor and let us focus on the empirically relevant case in which low-skilled individuals are more elastic, $|\epsilon_{L_u, A}| > |\epsilon_{L_s, A}|$.⁷ A permanent increase in the disutility of labor implies that both types of agents reduce their labor effort in the long run. Since low-skilled individuals are more elastic, they will reduce their labor effort more than high-skilled ones: the fraction of hours worked by the latter group, therefore, increases this pushing up labor productivity in the long run. The identification scheme will interpret this change in the composition of the labor force as a positive technology shock. Notice also that, under this restriction of the parameters, the long-run identification scheme induces a systematic downward bias in the estimation of the response of hours to technology shock. Similar considerations hold for the Solow residual accounting procedure. In Section 4 we will quantitatively assess the magnitude of this bias.

⁷A larger elasticity for low-skilled individual is consistent with the finding, discussed in Section 4 of this paper, that hours worked of low-skilled workers are more volatile than hours worked by high-skilled individuals.

2.3 Standard Approaches to Controlling for Input Heterogeneity

The fact that inputs heterogeneity complicates the measurement of technology is a well known problem in the growth accounting literature. Here we discuss the most widely accepted procedure that was developed by Jorgenson (1966). An alternative but closely related procedure due to Hansen (1993) is discussed in Appendix I. Central to these approaches is the approximation of the growth rate of L_t^e in terms of a weighted sum of the hours worked by different groups of individuals:

$$\Delta \log(L_t^e) \approx \sum_{j=1}^J a_{j,t} \Delta \log(L_{j,t}). \quad (11)$$

The procedures differ in the way the weights $\{a_{j,t}\}$ are computed. Jorgenson uses the following Tornqvist aggregator:

$$a_{j,t} = \frac{\nu_{j,t} + \nu_{j,t-1}}{2}, \quad (12)$$

and

$$\nu_{j,t} = \frac{w_{i,t} L_{i,t}}{\sum_j w_{j,t} L_{j,t}}. \quad (13)$$

As shown in Diewert (1976), this would be the right correction to make in the case that L_t^e is a deterministic homogeneous translog function of the J groups considered, $\log(L_t^e) = f(\log(\mathbf{L}_t))$, where \mathbf{L}_t is the vector of hours worked by the J groups.⁸ Using the properties of quadratic function (e.g., translog as defined in footnote 8), one obtains:

$$\begin{aligned} \Delta \log(L_t^e) &= f(\log(\mathbf{L}_t)) - f(\log(\mathbf{L}_{t-1})) \\ &= \frac{1}{2} [\nabla f(\log(\mathbf{L}_t)) + \nabla f(\log(\mathbf{L}_{t-1}))]' (\log(\mathbf{L}_t) - \log(\mathbf{L}_{t-1})), \end{aligned} \quad (15)$$

where the matrix $\nabla f(\log(\mathbf{L}_t))$ collects the partial derivatives of $f(\cdot)$. Under the additional assumption that prices equal marginal products at all points in time, the Jacobian $\nabla f(\log(\mathbf{L}_t))$

⁸ Defined by

$$\ln f(x) = \alpha_0 + \sum_{k=1}^K \alpha_k \ln x_k + \frac{1}{2} \sum_{m=1}^K \sum_{l=1}^K \gamma_{ml} \ln x_m \ln x_l, \quad (14)$$

where $\sum_{k=1}^K \alpha_k = 1$, $\gamma_{ml} = \gamma_{lm}$ and $\sum_{l=1}^K \gamma_{ml} = 0$ for $j = 1, 2, \dots, K$.

is equal to $\frac{w_{i,t}L_{i,t}}{\sum_j w_{j,t}L_{j,t}}$. Thus, equation 11 is exact for a homogeneous translog aggregator when the weights are Turnquist indexes of labor shares of different groups. All other functional forms, e.g., CES aggregator, will generate a bias.

A fundamental problem of this strategy arises when hours in efficiency units is not a deterministic aggregator of hours worked. An implicit assumption in this procedure is that the parameters of the aggregator have to be constant, making it for example difficult to explain movements in the skill premium. Thus even if the aggregator satisfies the functional form requirements at every point of time but parameters are changing over time, technology is measured with a bias. In order to make this point explicit, suppose that $\log(L_t^e) = f(\log(\mathbf{L}_t), \Theta_t)$, where Θ_t is a vector of time varying observable or unobservable factors and parameters. In this environment, one immediately verifies that equation 16 is an incorrect expansion for L_t^e as it neglects changes in Θ_t .

3 A New Method for Identifying Technology Shocks

In this section we show how neutral technology shocks can be identified in a wide class of models using standard filtering/smoothing techniques. The crucial insight is that equilibrium models share optimality conditions of the following form:

$$\mathbf{D}_t = Z_t + \mathbf{f}_t(\mathbf{D}_t), \tag{16}$$

where \mathbf{D}_t is a vector of observable variables and Z_t is the technology shocks. One example of the above restrictions is at the core of the Solow residual accounting, in which case \mathbf{D}_t include the logarithm of output and inputs. As another example, suppose that the aggregate production function in the economy is given by

$$Y_t = K_t^\alpha (e^{Z_t} L_{e,t})^{1-\alpha}, \tag{17}$$

where $L_{e,t} = F(L_{1,t}, \dots, L_{N,t}; \theta_t)$ is a general aggregator of hours worked by different types of agents. Then, assuming that factor prices equal marginal productivities one gets:⁹

$$\begin{aligned} & \log(W_{j,t}) + \frac{\alpha}{1-\alpha} \log(r_t) \\ &= \left[(1-\alpha) + \frac{\alpha}{1-\alpha} \log(\alpha) \right] + \log(Z_t) + \log \left(\frac{\partial F(L_{1,t}, \dots, L_{N,t}, \Theta_t)}{\partial L_{j,t}} \right) \end{aligned} \quad (18)$$

The structure of the system in equation (16) and (18) is that of a dynamic factor model where a set of observables depends on a common factor, Z_t , and on partially idiosyncratic factors. The method we propose does not require the researcher to specify a functional form for the aggregator $L_{e,t} = F(L_{1,t}, \dots, L_{N,t}; \theta_t)$. Instead, we treat both the technology series and the partial derivatives of $L_{e,t}$ as being unobserved and identify them using filtering/smoothing techniques.

The key identification challenge is that the common and idiosyncratic factors are contemporaneously related: a shock to technology is likely to move hours worked of different types of agents which in turn moves the factors $\log \left(\frac{\partial F(L_{1,t}, \dots, L_{N,t}, \Theta_t)}{\partial L_{j,t}} \right)$. In this section we present a set of sufficient conditions that allow one to identify the technology series from restrictions of the type presented in equation (16). For illustrative purpose, we will discuss our procedure within the context of equation (18).

First of all, in order to use standard tools adopted in the literature for the identification of state space model, we are going to make the following assumption concerning the DGP for Z_t and $\log \left(\frac{\partial F(L_{1,t}, \dots, L_{N,t}, \Theta_t)}{\partial L_{j,t}} \right)$:

Assumption 1.

Let $\mathbf{S}_t = \left[\left\{ \log \left(\frac{\partial F(L_{1,t}, \dots, L_{N,t}, \Theta_t)}{\partial L_{j,t}} \right) \right\}_{j=1}^k, \mathbf{Y}_t \right]'$ be a vector of dimension $n \times 1$, with \mathbf{Y}_t being a vector of dimension $(n-k) \times 1$ containing variables that are observed by the econometrician. We assume that \mathbf{S}_t has the following VAR(p) representation:

$$\mathbf{S}_t = \Phi_p(\mathbf{L})\mathbf{S}_t + \mathbf{e}_t, \quad (19)$$

⁹These conditions are going to be satisfied in models featuring constant returns to scale and perfect competition. A similar version for the system in ((18)) holds also in models with non-constant return to scale and time varying mark-ups over factor prices. See the Appendix for details.

where $\Phi_p(\mathbf{L})$ is a polynomial of order p in the lag operator \mathbf{L} . We also assume that Z_t follows the first order process:

$$Z_t = \rho_z Z_{t-1} + e_{z,t}. \quad (20)$$

As Appendix III.1 shows, this assumption is satisfied in a wide class of models whenever $p \rightarrow \infty$. Data limitation necessitate that we rely on the finite dimensional approximation to the $VAR(\infty)$.

Given Assumption 1, we can write down our state space model as follows:

$$\mathbf{D}_t = \begin{bmatrix} \mathbf{1} & \mathbf{I}_J & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_K \end{bmatrix} \begin{bmatrix} Z_t \\ \mathbf{S}_t \end{bmatrix} = \mathbf{B} \begin{bmatrix} Z_t \\ \mathbf{S}_t \end{bmatrix} \quad (21)$$

$$\begin{bmatrix} Z_t \\ \mathbf{S}_t \end{bmatrix} = \underbrace{\begin{bmatrix} \rho_z^o & \mathbf{0} \\ \Phi_z^o & \Phi_S^o \end{bmatrix}}_{\Phi^o} \begin{bmatrix} Z_{t-1} \\ \mathbf{S}_{t-1} \end{bmatrix} + \begin{bmatrix} e_{z,t} \\ \mathbf{e}_t \end{bmatrix} \quad \begin{bmatrix} e_{z,t} \\ \mathbf{e}_t \end{bmatrix} \sim \mathcal{WN}(\mathbf{0}, \Sigma^o) \quad (22)$$

where \mathbf{D}_t is a vector collecting $W_{j,t} + \frac{\alpha}{1-\alpha} R_t$ for every observed group j and \mathbf{Y}_t .

Assumption 2.

$$\mathbf{E}[\mathbf{D}_t | \mathbf{D}^{t-1}] \neq \mathbf{E}[\mathbf{D}_t | \log(Z_{t-1}), \mathbf{S}^{t-1}].$$

Assumption 2 ensures that the state space considered is minimal.

Lemma 1. *Let Assumption 2 hold. Then, the state space model in (21) and (22) is minimal.*

Proof. See Appendix III.2. □

Given minimality, lack of identification in our state space model is known to be represented by linear transformations of the state vector through an invertible matrix \mathbf{T} . Thus, denoting by an “ $\hat{\cdot}$ ” an observationally equivalent structure, one must have that:

$$\mathbf{T} \begin{bmatrix} \rho_z & \mathbf{0}' \\ \Phi_z & \Phi \end{bmatrix} \mathbf{T}^{-1} = \begin{bmatrix} \hat{\rho}_z & \mathbf{0}' \\ \hat{\Phi}_z & \hat{\Phi} \end{bmatrix} \quad (23)$$

$$\mathbf{B}\mathbf{T} = \mathbf{B} \quad (24)$$

$$\mathbf{T}\Sigma^{1/2} = \hat{\Sigma}^{1/2} \quad (25)$$

The assumptions made so far restrict the set of admissible \mathbf{T} matrices, but they are still not sufficient to guarantee that the state space system is identified. We impose one of the following two assumptions:

Assumption 3. *The vector of innovations \mathbf{e}_t is related to the structural innovations as follows:*

$$\mathbf{e}_t = \mathbf{A}\eta_t \quad \eta_t \sim \mathcal{WN}(\mathbf{0}, \mathbf{I})$$

Assumption 4. *There exists a vector of instruments \mathbf{x}_t , orthogonal to $e_{z,t}$ and such that $\text{cov}\left(\mathbf{x}_t, \left\{\log\left(\frac{\partial \mathbf{F}(\mathbf{L}_{1,t}, \dots, \mathbf{L}_{N,t}, \theta_t)}{\partial \mathbf{L}_{j,t}}\right)\right\}_{j=1}^k\right)$ is full rank.*

Proposition 2. *Suppose that **Assumption 3** holds. Then, $\theta^o = (\Phi^o, \Sigma^o)$ is locally identified from the probability density of $\{\mathbf{D}_t\}$.*

Proposition 3. *Suppose that **Assumption 4** holds. Then, up to a scale and a sign normalization, $\theta^o = (\Phi^o, \Sigma^o)$ is globally identified from the probability density of $\{\mathbf{D}_t, \mathbf{x}_t\}$.*

Proposition 4. *Suppose that **Assumption 3** and **Assumption 4** hold. Then, $\theta^o = (\Phi^o, \Sigma^o)$ is globally identified from the probability density of $\{\mathbf{D}_t, \mathbf{x}_t\}$.*

Let us illustrate the idea of the proof in the case that we do not have observed states (the state vector does not include \mathbf{Y}_t). Given equation 24 and the expression for \mathbf{B} , the matrix \mathbf{T} has to satisfy:

$$\mathbf{I} + \begin{bmatrix} \kappa_1 & \kappa_2 & \dots & \kappa_i & \dots & \kappa_{k+1} \\ \kappa_1 & \kappa_2 & \dots & \kappa_i & \dots & \kappa_{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \kappa_1 & \kappa_2 & \dots & \kappa_i & \dots & \kappa_{k+1} \\ \kappa_1 & \kappa_2 & \dots & \kappa_i & \dots & \kappa_{k+1} \end{bmatrix}.$$

Notice that since $[Z_t, \mathbf{S}_t]'$ and $\mathbf{T}[Z_t, \mathbf{S}_t]'$ are not distinguishable, one has that $\log(\hat{Z}_t) = \mathbf{e}_1 \mathbf{T}[Z_t, \mathbf{S}_t]'$ is not distinguishable from $\log(Z_t)$. Given the restrictions we have so far on \mathbf{T} one has that:

$$\log(\hat{Z}_t) = \log(Z_t)(1 + \kappa_1) + \kappa_2 \log(F_{L_1,t}) + \kappa_3 \log(F_{L_2,t}) + \dots + \kappa_{k+1} \log(F_{L_k,t}). \quad (26)$$

By Assumption 4 we have that $Cov(x_t^i, \log(Z_t)) = 0$ for every instrument instrument. Imposing this condition in our system, we must also have that $Cov(x_t^i, \log(\hat{Z}_t)) = 0$. Thus, by equation (26), one has:

$$\mathbf{Cov} \left(\mathbf{x}_t, \left\{ \log \left(\frac{\partial F(L_{1,t}, \dots, L_{N,t}, \Theta_t)}{\partial L_{j,t}} \right) \right\}_{j=1}^k \right) \begin{bmatrix} \kappa_2 \\ \kappa_3 \\ \dots \\ \kappa_{k+1} \end{bmatrix} = \mathbf{0}.$$

Since the covariance matrix has full rank by Assumption 4, the above restrictions imply $\kappa_j = 0$ for all $j > 1$. Therefore, the matrix T has the following form:

$$\begin{bmatrix} 1 + \kappa_1 & 0 & \dots & 0 & \dots & 0 \\ \kappa_1 & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \kappa_1 & 0 & \dots & 0 & \dots & 0 \\ \kappa_1 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}.$$

In this case we have that

$$\log(\hat{F}_{L_k,t}) = \log(Z_t)\kappa_1 + \log(F_{L_k,t}). \quad (27)$$

Under the condition that $\sum_{m=1}^k \phi_{jm} \neq \rho_z$ for at least one j , we have that $\kappa_1 = 0$ through the restrictions that the first column of the transition matrix is zero as is implied by Assumption 3. One can easily verify that if Assumption 2 holds, then we can always find such a j . This implies that $\mathbf{T} = \mathbf{I}$, that is the state space is identified.

4 Monte Carlo Simulations using a Calibrated Model with Heterogeneous Labor

In this section we calibrate the standard RBC model with two types of workers, simulate data from it and study the relation between identified “technology” shocks and the true structural disturbances in the model economy. In particular, we consider the small sample performance of the method for identifying technology shocks proposed in this paper and contrast it with the performance of the standard methods.

4.1 Model

We consider an RBC model with worker heterogeneity described in Section 2. In particular, we assume that agents of type $j = \{u, s\}$ (unskilled and skilled) value consumption and leisure according to type-dependent utility functions

$$U_j(c_t, h_t) = \log(c_t) - e^{A_t} b_j \frac{h_t^{1+\nu_j^{-1}}}{1 + \nu_j^{-1}}. \quad (28)$$

Firms have access to the production function

$$Y_t = K_t^\alpha (e^{Z_t} L_t^e)^{1-\alpha}, \quad (29)$$

where L_t^e , the effective labor inputs, is an aggregator of low and high-skilled labor inputs

$$L_t^e = L_{h,t}^{\phi_t} L_{l,t}^{1-\phi_t}.$$

Technology and preference shocks are assumed to follow the AR(1) processes:

$$\begin{aligned} Z_t &= \rho_z Z_{t-1} + \sigma_z \varepsilon_{z,t}, \\ A_t &= \rho_a A_{t-1} + \sigma_a \varepsilon_{a,t}, \end{aligned}$$

while $\phi_t \in [0, 1]$ follows the law of motion:

$$\phi_t = \frac{e^{\tilde{\phi}_t}}{1 + e^{\tilde{\phi}_t}} \quad \tilde{\phi}_t = (1 - \rho_\phi) \mu_\phi + \rho_\phi \tilde{\phi}_{t-1} + \sigma_\phi \varepsilon_{\phi,t}$$

The innovation $\varepsilon_t = [\varepsilon_{z,t}, \varepsilon_{a,t}, \varepsilon_{\phi,t}]'$ are i.i.d. standard normal random variates.

Appendix II discusses the equilibrium relations as well as the solution method adopted.

4.2 Calibration

The vector of structural parameters of our model is given by:

$$\theta = \underbrace{[\beta, \delta, \alpha, b_u, b_s, l, \mu_\phi]}_{\theta_1} \underbrace{[\rho_\phi, \sigma_\phi]}_{\theta_2} \underbrace{[\bar{v}, x_l, \rho_z, \sigma_z, \rho_a, \sigma_a]}_{\theta_3}.$$

Model period is one quarter. We use quarterly postwar data on the US economy in order to estimate the vector θ . The parameters in θ_1 are pinned down using long run average for selected

time series. In particular, the parameters β , α and δ are chosen so that, in a deterministic steady state of the model, the real interest rate, the depreciation rate of capital and a labor income share are respectively 1%, 2.5% and 66%, values that are common in the business cycle literature. The parameters b_u , b_s , l and μ_ϕ are chosen so that the model matches 22.6 weekly hours worked by low-skilled individual, 31.8 weekly hours worked by high-skilled individuals, a fraction of low-skilled individuals over total population of 0.774 and a skill premium equal to 1.7. These numbers are calculated using CPS quarterly data (1976-2006) on wages and hours worked by education level.¹⁰

The parameters in θ_2 are estimated using quarterly observations on wages and hours worked by education group. In particular, the model implies that:

$$\tilde{\phi}_t = \log \left(\frac{W_{s,t} H_{s,t}}{W_{u,t} H_{u,t}} \right).$$

Given our definition of low and high skilled individuals, we obtain a quarterly series on wages and hours worked from CPS data and estimate an AR(1) process for hp-filtered ($\lambda = 1600$) $\tilde{\phi}_t$ using OLS.

The remaining parameters in θ_3 are estimated via a standard Simulated Method of Moments (SMM) algorithm. In particular, let \mathbf{m} be a vector of moments for selected time series computed using US time series of length \mathbf{T} and let $\mathbf{m}_{\mathbf{T}}(\theta)$ be their model counterpart when the vector of structural parameter is $\theta = \{\theta_1^*, \theta_2^*, \theta_3\}$. The estimator for θ_3 solves:

$$\min_{\theta_3} \quad [\mathbf{m}_{\mathbf{T}} - \hat{\mathbf{m}}(\{\theta_1^*, \theta_2^*, \theta_3\})]' \mathbf{W}_{\mathbf{T}} [\mathbf{m}_{\mathbf{T}} - \hat{\mathbf{m}}(\{\theta_1^*, \theta_2^*, \theta_3\})],$$

where $\mathbf{W}_{\mathbf{T}}$ is a diagonal matrix whose nonzero elements are the inverse of the variance of the corresponding moment.

As for the empirical moments, we include in the vector \mathbf{m} standard measures of cyclical volatility and comovement for hp-filtered ($\lambda = 1600$) quarterly US data.¹¹ The time series adopted are output for the business sector, private non-durable consumption, private non-residential investment, total hours worked in the business sector, total hours of low and high

¹⁰We define high-skilled individuals as those possessing college education and low-skilled individuals as those with no college education. See Appendix B.

¹¹All time series are in logs. As measures of volatility we include the standard deviation while for comovement we include the correlation coefficient.

skilled individuals in the business sector and labor productivity. Moreover, in view of our previous discussion, we include in \mathbf{m}_T the empirical IRFs to a “long run” shock identified using Gali (1999) methodology on a bivariate VAR(4) on the growth rate of labor’s productivity and total hours worked for the US economy. The horizon of the IRFs is 5.

Simulated moments are calculated using a Monte Carlo procedure. In particular, for each θ , we solve the model using first order perturbation, simulate a realization of length $T = 250$ for the model’s counterparts of the above time series and we calculate the vector $\hat{\mathbf{m}}_T(\theta)$ using the same procedure on the model generated data. We repeat this procedure $M = 300$ times, each time changing the seed used in the simulation. We then take the (component wise) median of $\hat{\mathbf{m}}(\theta)$ across the Monte Carlo replications.

We apply standard numerical minimization routines to estimate θ_3 . Since the minimization algorithm was pushing preference shocks to their non-stationary region, we have decided to impose $\rho_a = 1$ and to reformulate the model for normalized (stationary) variables. This amounts to scale variable i in the model by the $e^{\Lambda_{i,a}A_t}$ with $\Lambda_{i,a}$ being the long run effect of a preference shock on variable i . Details on this normalization procedure and on the method adopted to solve the log-linearized model can be found in Appendix II. Table 1 summarizes the procedure used for the calibration of our model and reports numerical values for the structural parameters along with robust standard errors. Table 2 and Figure 2 report the fit our model in terms of the targets in the SMM estimation.

As we can see from Figure 1, preference shocks have permanent effects on all model variables. In particular, following an increase in A_t , both low-skilled and high-skilled individuals reduce their labor supply, this implying that aggregate hours falls permanently. The reduction in aggregate hours, though, is due mainly to the low-skilled individuals since their labor supply is more elastic. The decline in total output, therefore, is not dramatic since aggregate hours fall mainly due to workers with low productivity: as a result, output per worker increases permanently following a positive innovation to A_t . A positive innovation to technology has transitory effects on model’s variables: hours worked, total output and productivity increase on impact and decay slowly toward zero.

To evaluate the performance of the model in matching the calibration targets in Table 2 we report the measures of cyclical volatility, persistence and comovement collected in the

vectors $\hat{\mathbf{m}}(\theta^*)$ and \mathbf{m}_T . The model is consistent along many dimensions with the behavior of aggregate time series at business cycles' frequencies. We are able to match volatilities and the correlation of the time series considered with real GDP. Conditional on a preference shock, labor productivity is negatively correlated with hours worked and with output. Since preference shocks drive most of the fluctuation in hours, our model predicts a negative correlation between worked hours and productivity, a statistic that is often difficult to match in RBC models. Despite that, as the variance decomposition in Table 2 shows, technology shocks still explain most of the variation in output, investment and productivity at business cycles' frequencies in the model.

Figure 2 plots IRFs retrieved using long run restrictions on a VAR(4) for the growth rate of labor's productivity and hours worked. In this graph, the blue (solid) line describe the IRFs retrieved from data simulated from our economy at the calibration considered in Table 1, the red (dashed) line refers to the empirical IRFs and black lines reports 95% bootstrap bands for the empirical IRFs. As we can see from the figure, a researcher that applies the Blanchard and Quah (1989) methodology on data simulated from our model would reach conclusions similar to those reported in Galí (1999) since the blue line tracks closely its empirical counterpart.

4.3 Identifying Technology Shocks in Model-Generated Data

We simulate 30 samples of 250 quarters each from the calibrated model. For each sample we estimate technology shocks using

- i) Solow residual with Jorgenson's correction for labor composition effects.¹²
- ii) SVAR with long-run restrictions as in Galí (1999).
- iii) The Unobserved Component model (UC) proposed in this paper. As instruments we use noisy signals over $\varepsilon_{z,t}$ and $\varepsilon_{\phi,t}$ with signal to noise ratio of 50%, 25%, and 10%.

For each Monte Carlo replication m , we estimate the following linear regression:

$$\varepsilon_{j,t}^{true} = \alpha + \beta \varepsilon_{z,t}^{identified} + \eta_t \quad \text{J} \in \{z, a, \phi\},$$

¹²When calculating the Solow residual we use the true parameter α rather than estimating it.

and we store the $R^2(m)$. We then average across the Monte Carlo replications and report the results in Table 3. Clearly, a method that recovers technology innovations perfectly (up to a scale factor) would obtain an R^2 of 1 in the technology equation, and zero in the other equations.

We can verify from Table 3 that, consistent with the argument of Section 2, the Long-Run shock and the Solow residuals have little structural interpretation in our model. Both series represent a combinations of shocks to technology, preferences, and the skill premium, with non trivial weights placed on the latter two. The Long Run shocks captures 25% of the variation in preference shocks, while shocks to the skill premium are misinterpreted as an indicator of technological change in the growth accounting procedure. This is true even when we apply Jorgenson correction to account for labor compositional effect. As we can verify from the Table, the Solow residual corrected for labor composition accounts for 20% of the variation in the skill premium shock. The evidence of Table 3 suggests that when data are generated from an empirically plausible version of the RBC model with workers heterogeneity the commonly used practice of backing-out the technology series either as a residual of an aggregate production function or as the shock affecting labor productivity in the long run does not produce reliable results. Changes in the composition of the labor force induced by exogenous demand shocks or by changes in the relative productivities of workers are misinterpreted by the two procedures as movement in technology.

Table 3 also illustrates that the method proposed in this paper performs quite well in identifying the true technology shocks. In particular, the identified technology shocks are nearly perfectly correlated with the true technology shocks and are uncorrelated with the preference shocks and the shocks to the relative productivity of skilled workers. Clearly, more informative signals result in better inference: the R^2 in the technology equation ranges from 98%-91% when using instruments whose signal to noise ratio that ranges from 50%-10%.

Figure 3-5 give a graphical intuition of the results in Table 3. In this figure we plot for a particular realization of our model a scatter plot of the structural shocks against the technology shocks identified using SVAR procedure with long-run restriction, and Solow residual with Jorgenson's correction, and the unobserved component model discussed in this paper. The technology shocks identified with the existing procedures are clearly correlated with pref-

erence shocks and the shocks to the skill premium while inference with the proposed method appears robust in this dimension.

Figure 6 reports a response surface for the relation between true and identified technology shocks when using the procedure described in Section 3. In particular, we consider different combinations of the signal to noise ratio and plot this measure of accuracy against these different combinations. The results suggests that the procedure is robust to using instruments that are informative at different degrees.

In Table 4 we show how each method performs when estimating the IRFs of hours worked to the identified technology shocks. In particular, we consider the following measure of bias:

$$\mathbf{Bias}_j = \frac{\left| \frac{\partial L_{t+j}}{\partial \varepsilon_{z,t}^{\text{true}}} - \mathbb{E} \left[\frac{\partial L_{t+j}}{\partial \varepsilon_t^{\text{identified}}} \right] \right|}{\left| \frac{\partial L_{t+j}}{\partial \varepsilon_{z,t}^{\text{true}}} \right|} \times 100$$

From the Table we can verify that inference over the response of hours worked to a technology shocks is imprecise. In particular, the average bias is 150% for the Long Run shock while it is 130% for the Solow residual. This large bias in the calculation of the response of hours to technology shocks are due to the combination of two factors: the two procedures capture a substantial ammount of other shocks in the economy; these non-technology shocks drive a substantial fraction of hours worked in our model at business cycle frequencies.

5 Technology Shocks in the U.S. Data

In this section we apply our method to identify technology shocks in U.S. data.

5.1 Instruments

We plan to use balanced growth restrictions. In fact, if preferences and technology are consistent with balanced growth, one gets that:

- i) Neutral technology shocks can not induce a trend in hours per capita;
- ii) Neutral technology shocks can not induce a trend in relative wages of different groups (e.g. skill premium);

We can then use low frequency information in wages and hours worked of different types of workers in order to construct our instruments for the partial derivatives of the aggregator $L_{e,t}$.

5.2 Findings

In the next draft we will describe and analyze the sequence of identified shocks and document its co-movement with other economic aggregates.

6 Conclusion

In this paper we have shown that the standard methods for identifying the technology shocks in the data are biased in models with heterogeneous inputs. In particular, presence of worker heterogeneity invalidates the key identification assumption in Blanchard and Quah (1989) and Gali (1999) because not only technology, but virtually all persistent shocks have a long run effect on productivity in such models. The identification of technology shocks using the production function estimation is also biased if the effects of worker heterogeneity are not explicitly accounted for. In particular, if less productive workers also have a more elastic labor supply, Solow residuals will be negatively correlated with labor input in response to preference shocks. The standard procedures used to correct for labor composition rely on strong assumptions that are unlikely to be satisfied in the data.

We proposed a new method to identify technology shocks in the data. We interpret the technology shock as an unobserved state and using the structure of the model we show that this state can be identified using filtering techniques. Since we do not treat the technology shock as a residual our method does not require to specify an explicit function that aggregates heterogeneous labor inputs. Moreover, we do not even require the parameters of this function to be invariant over time.

We evaluate the small sample properties of our method using an estimated RBC model that features two imperfectly substitutable types of labor (skilled and unskilled) and persistent demand and supply shocks. We find that the model is quantitatively consistent with unconditional RBC statistics and conditional correlations identified by a structural vector

autoregression with long run restrictions. Most of the cyclical fluctuations in output and productivity in the model are driven by technology shocks which the identification strategy based on long run restrictions misinterprets as largely representing demand disturbances. Most of the fluctuations in hours, on the other hand, are accounted for by shocks to preferences. We find that our method performs quite well in identifying technology shocks in the model generated data as compared to the existing methods.

Table 1: **Calibrated Parameters**

Parameter	Value	Data used to Identify Parameters
α	0.330	Labor Income Share
δ	0.025	Depreciation of Capital Stock
β	0.995	Real Interest Rate
h_s^*	0.318	Weekly Hours per Individual (College)
h_u^*	0.226	Weekly Hours per Individual (no College)
u	0.774	% of Individuals without College
μ_ϕ	-0.140	Skill Premium
ρ_ϕ	0.522	Estimated by OLS
$\sigma_\phi \times 100$	3.000	Estimated by OLS
$\bar{\nu}$	1.500	Fixed
ν_s	0.526	Estimated by SMM
ρ_a	1.000	Estimated by SMM
ρ_z	0.950	Estimated by SMM
$\sigma_z \times 100$	1.140	Estimated by SMM
$\sigma_a \times 100$	1.820	Estimated by SMM

Note: The time horizon is quarterly. In the SMM algorithm we match (1) standard deviation, first order autocorrelation and correlation with GDP of the following h-p filtered quarterly time series: real GDP, labor productivity, hours worked (total and by skill group), consumption and investment. (2) The impulse response functions of a VAR(4) on labor productivity and hours worked (1948-1994). Shocks are identified imposing long-run restrictions on labor productivity as in Gali (1999).

Table 2: **Estimation Targets: Business Cycle Statistics**

		Y_t	I_t	C_t	L_t	$\frac{Y_t}{L_t}$	
St. Deviation (%)	Data	2.07	5.13	1.16	1.85	1.08	
	Model	1.81	5.27	0.67	1.78	0.97	
Corr. with Y_t	Data	1.00	0.78	0.75	0.86	0.45	
	Model	1	0.98	0.82	0.85	0.30	
% St. Dev., L_s	Data	1.43			% St. Dev., L_u	Data	2.15
	Model	1.04				Model	2.28
Corr. L_t and $\frac{Y_t}{L_t}$	Data	-0.07			Corr. L_s and L_u	Data	0.51
	Model	-0.25				Model	0.56

Note: All series are quarterly (see Appendix x for data definition). Prior to compute all relevant statistics, variables have been transformed through natural logarithms and hp-filtered ($\lambda = 1600$). We perform the same operations on data simulated from our model (each data series has length 250) and report the median over the 300 Monte Carlo replications.

Table 3: **True Shocks vs. Identified Technology Shocks**

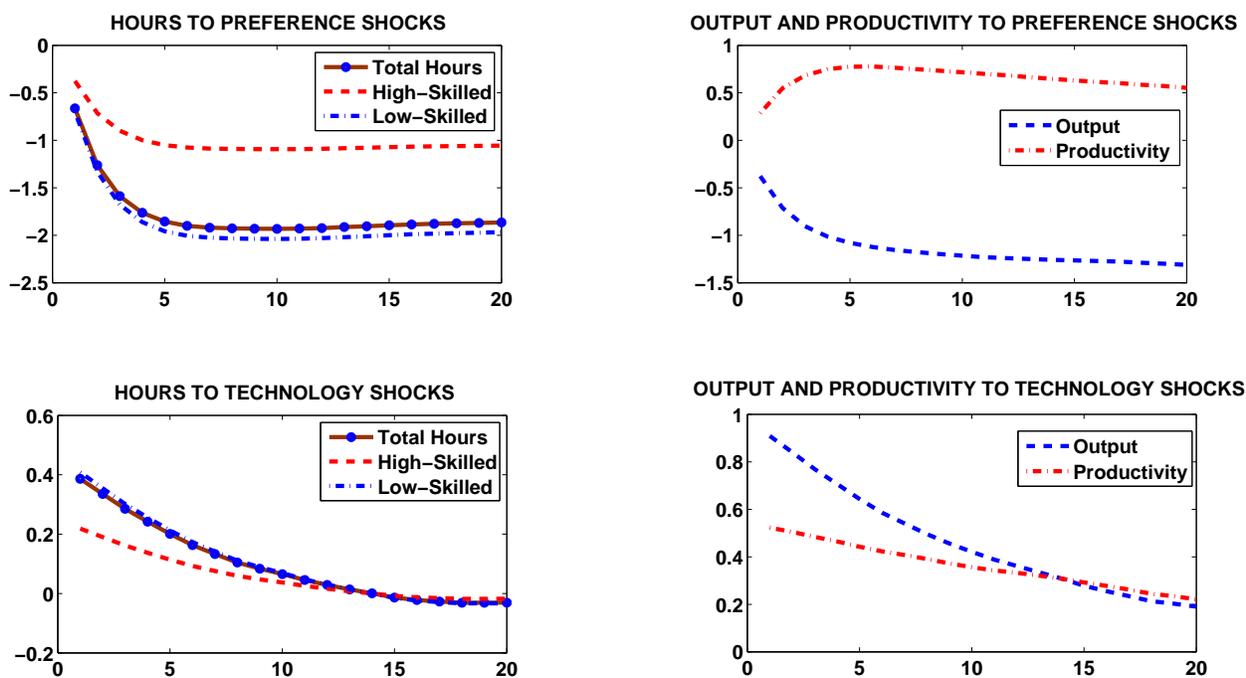
	$\mathbb{E}[R^2(\eta_{z,t})]$	$\mathbb{E}[R^2(\eta_{\phi,t})]$	$\mathbb{E}[R^2(\eta_{a,t})]$
Long-Run SVAR	0.70	0.03	0.25
Solow (Corrected)	0.74	0.20	0.01
UC (50%)	0.98	0.01	0.01
UC (25%)	0.96	0.01	0.02
UC (10%)	0.91	0.02	0.06

Note: Results are based on a Monte Carlo studies with 30 replications. The Solow residual is calculated using Equation (4) and applying Jorgenson correction for labor composition effects. The Long-Run shock is calculated applying the procedure described in Gali (1999) on the simulated data. We estimate the Unobserved Component model and back out the realization of the technology innovations as discussed in Section 3.1.

Table 4: **Bias in the estimated response of hours to technology**

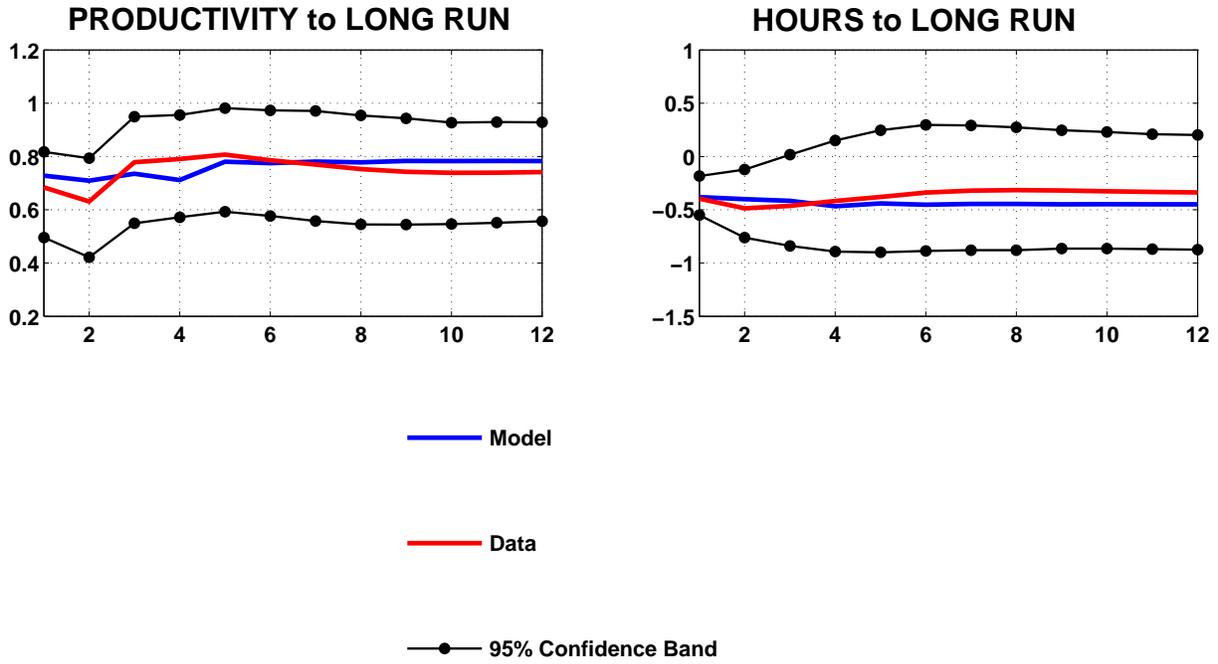
Horizon (j)	0	1	2	3	4
Long-Run shock	154%	155%	175%	184%	185%
Solow Residual (Corrected)	188%	158%	157%	148%	148%
UC	23%	10%	13%	12%	1%

Figure 1: Impulse Response Functions to Preference and Technology Shocks



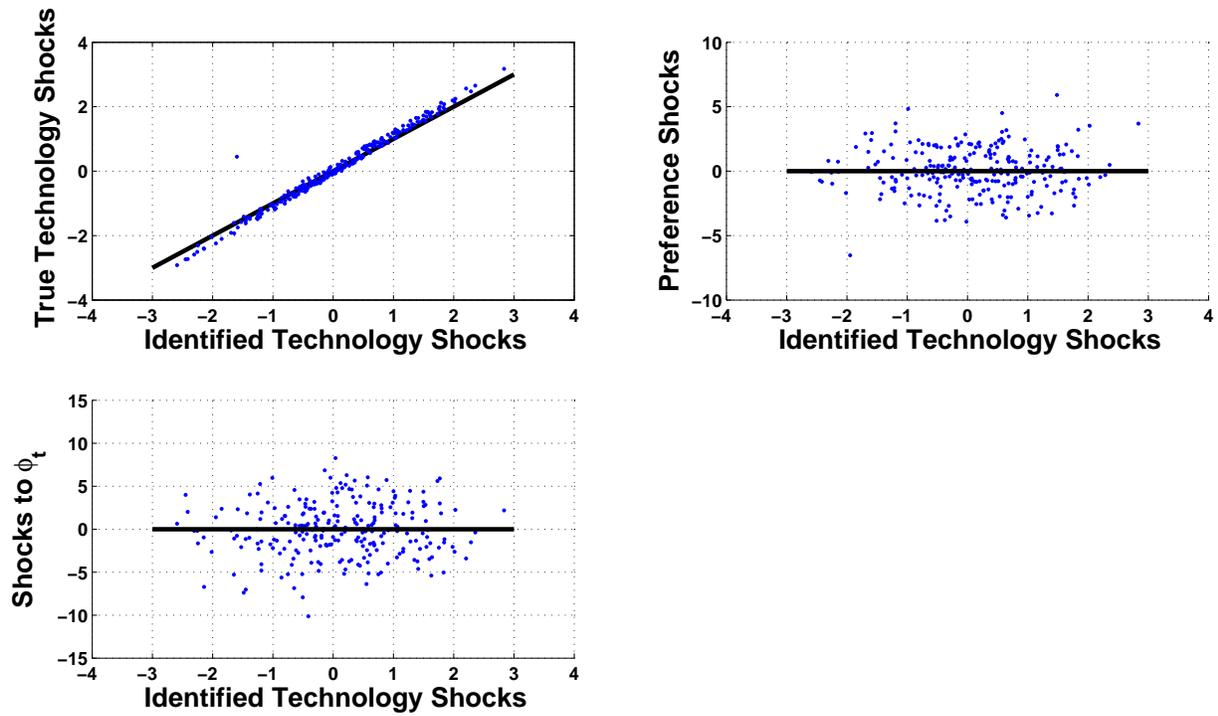
Note: Response of selected model variables (%) to a one standard deviation increase in preference and technology innovations.

Figure 2: Impulse Response Functions



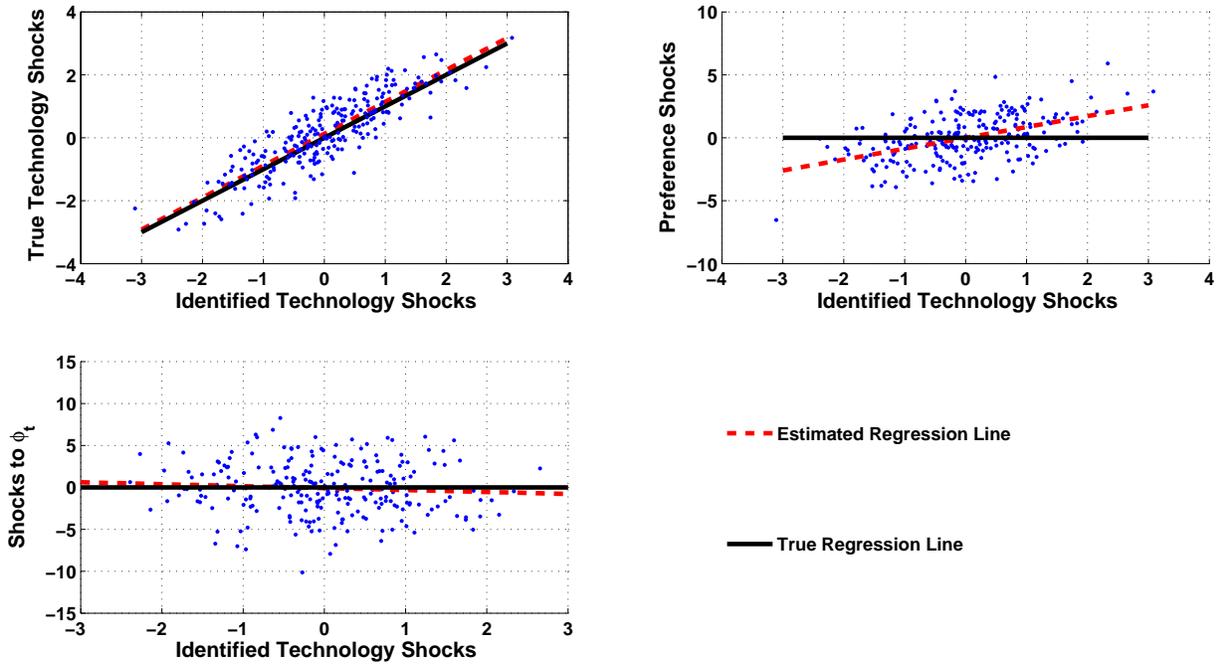
Note: Note: The model IRFs (blue line) are calculated by applying Gali (1999) methodology on data simulated from our model. The data IRFs (red line) replicates Gali (1999) findings using US quarterly data (1948-1994) on real GDP and non-farm business hours worked. Solid-bullet lines report 95% bootstrap confidence band for the empirical IRFs.

Figure 3: Structural Shocks vs Technology Shocks Identified using the Method Proposed in this Paper



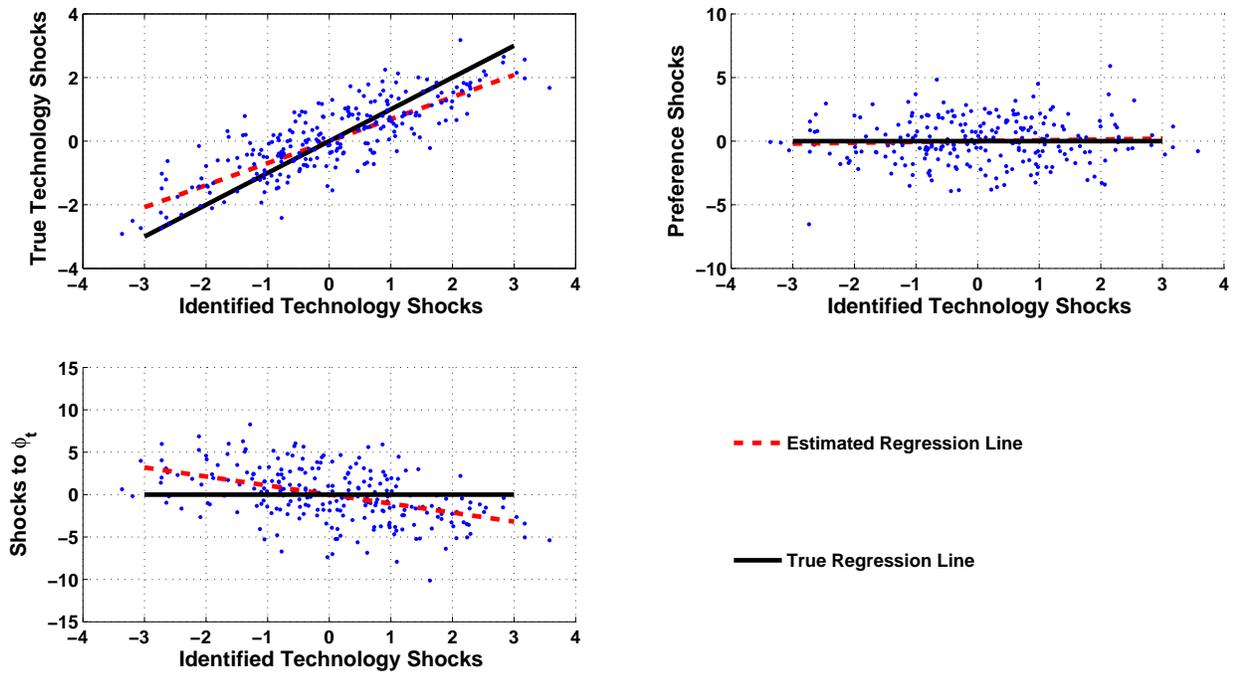
Note: .

Figure 4: Structural Shocks vs Technology Shocks Identified using VAR with Long-Run Restrictions



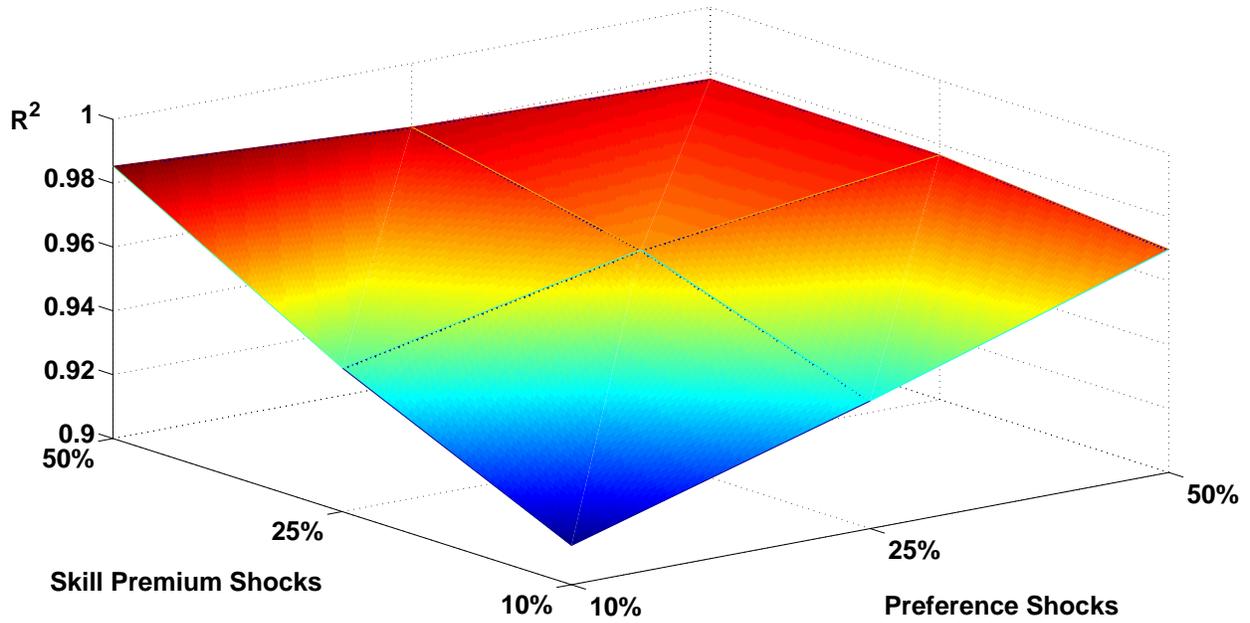
Note: .

Figure 5: Structural Shocks vs Technology Shocks Identified as Solow Residuals



Note: .

Figure 6: Response Surface



Note: .

References

- ALEXOPOULOS, M. (2010): “Read All About It! What Happens after a Technology Shock?,” *American Economic Review*, Forthcoming.
- BASU, S., J. G. FERNALD, AND M. S. KIMBALL (2006): “Are Technology Improvements Contractionary?,” *The American Economic Review*, 96(5), 1418–1448.
- BENCIVENGA, V. R. (1992): “An Econometric Study of Hours and Output Variation with Preference Shocks,” *International Economic Review*, 33(2), 449–471.
- BENHABIB, J., R. ROGERSON, AND R. WRIGHT (1991): “Homework in Macroeconomics: Household Production and Aggregate Fluctuations,” *Journal of Political Economy*, 99(6), 1166–1187.
- BLANCHARD, O. J., AND D. QUAH (1989): “The Dynamic Effects of Aggregate Demand and Supply Disturbances,” *American Economic Review*, 79(4), 655–673.
- CHARI, V., P. J. KEHOE, AND E. R. MCGRATTAN (2008): “Are structural VARs with long-run restrictions useful in developing business cycle theory?,” *Journal of Monetary Economics*, 55(8), 1337–1352.
- CHRISTIANO, L. J., AND M. EICHENBAUM (1988): “Is Theory Really Ahead of Measurement? Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations,” Working Paper No. 2700, National Bureau of Economic Research.
- CHRISTIANO, L. J., M. EICHENBAUM, AND R. J. VIGFUSSON (2006): “Alternative Procedures for Estimating Vector Autoregressions Identified with Long-run Restrictions,” *Journal of the European Economic Association*, 4(2-3), 475–483.
- DIEWERT, W. E. (1976): “Exact and Superlative Index Numbers,” *Journal of Econometrics*, 4(2), 115–145.
- FAUST, J., AND E. LEEPER (1997): “When Do Long-Run Identifying Restrictions Give Reliable Results?,” *Journal of Business and Economic Statistics*, 15(3), 345–353.

- FERNANDEZ-VILLAVERDE, J., J. F. RUBIO-RAMIREZ, T. J. SARGENT, AND M. W. WATSON (2007): “ABCs (and Ds) of Understanding VARs,” *American Economic Review*, 97(3), 1021–1026.
- FRANCIS, N., AND V. A. RAMEY (2005): “Is the Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited,” *Journal of Monetary Economics*, 52(8), 1379–1399.
- GALI, J. (1999): “Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?,” *American Economic Review*, 89(1), 249–271.
- (2004): “On the Role of Technology Shock as a Source of Business Cycles: Some New Evidence,” *Journal of the European Economic Association*, 2(1-2), 372–380.
- HANSEN, G. (1993): “The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked,” *Journal of Applied Econometrics*, 8(1), 71–80.
- JORGENSON, D. W. (1966): “The Embodiment Hypothesis,” *Journal of Political Economy*, 74(1), 1–17.
- KATZ, L. F., AND K. MURPHY (1992): “Changes in Relative Wages: Supply and Demand Factors,” *Quarterly Journal of Economics*, 107, 35–78.
- KING, R. G., C. I. PLOSSER, J. H. STOCK, AND M. W. WATSON (1991): “Stochastic Trends and Economic Fluctuations,” *American Economic Review*, 81, 819–840.
- KYDLAND, F. E., AND E. C. PRESCOTT (1982): “Time to Build and Aggregate Fluctuations,” *Econometrica*, 50, 1345–1370.
- LONG, J. B. J., AND C. I. PLOSSER (1983): “Real Business Cycles,” *Journal of Political Economy*, 91(1), 36–69.
- PRESCOTT, E. (2004): “Why Do Americans Work So Much More Than Europeans,” *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1), 2–13.
- SHEA, J. (1998): “What Do Technology Shocks Do?,” in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke, and J. Rotemberg, vol. 13. The MIT Press.

SOLOW, R. M. (1957): “Technical Change and the Aggregate Production Function,” *The Review of Economics and Statistics*, 39(3), 312–320.

UHLIG, H. (2004): “Do Technology Shocks Lead to a Fall in Total Hours Worked?,” *Journal of the European Economic Association*, 2(2-3), 361–371.

URIBE, M., AND S. SCHMITT-GROHE (2004): “Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function,” *Journal of Economic Dynamics and Control*, 28(4), 755–775.

APPENDICES

I Hansens' Procedure for Controlling for Input Heterogeneity

Hansen (1993) measures the efficiency units of labor as

$$\sum_i \alpha_i L_{i,t}, \tag{A1}$$

where α_i is the constant weight of group i . The weights α_i are the average hourly earnings

$$\alpha_i = \frac{HE_i}{HE}, \tag{A2}$$

where HE_i is average hourly earnings for group i and HE is average hourly earnings.

We first compute a log-linear approximation of $\log(\sum_i \alpha_i L_{i,t})$ with respect to $\log(L_{i,t})$:

$$\log\left(\sum_i \alpha_i L_{i,t}\right) \approx \sum_i \frac{\alpha_i \bar{L}_i}{\sum_j \alpha_j \bar{L}_j} \log(L_{i,t}), \tag{A3}$$

where \bar{L}_i is the average labor supply of group i . In addition to this approximation, a second difference between Hansen and Jorgensen is that they use different coefficients. Jorgensens uses $\nu_{j,t}$, an average of two adjacent periods whereas Hansen uses

$$\frac{\alpha_i \bar{L}_i}{\sum_j \alpha_j \bar{L}_j}, \tag{A4}$$

a time average for the full sample. This means the second bias in the measurement due to differences in computing averages of wages equals

$$\nu_{j,t} - \frac{\alpha_i \bar{L}_i}{\sum_j \alpha_j \bar{L}_j}. \tag{A5}$$

After these approximations, Hansen measurement is equal to Jorgenson and thus is unbiased if and only if the aggregator is a homogeneous translog function (with constant coefficients).

II Model Solution and its State-Space Representation

The decision problem of the Social Planner is:

$$\max_{\{c_{j,t}, h_{j,t}\}_{j \in \{H,L\}, K_{t+1}\}_{t=0}^{\infty}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \sum_{j \in \{H,L\}} \tilde{\psi}_j \beta^t \left[\log(c_{j,t}) - e^{A_t} b_j \frac{h_{j,t}^{1+\nu^{-1}}}{1+\nu^{-1}} \right] \right\}$$

$$(1-m)c_{h,t} + mc_{l,t} + K_{t+1} - (1-\delta)K_t \leq K_t^\alpha \left\{ e^{Z_t} [(1-m)h_{h,t}]^\phi [mh_{l,t}]^{1-\phi} \right\}^{1-\alpha} \quad (\text{A6})$$

$$h_{j,t} \in [0, 1] \quad c_{j,t} \geq 0 \quad K_t \geq 0 \quad \forall t \geq 0 \quad \forall j \in \{H, L\} \quad (\text{A7})$$

We denote by $\tilde{\psi}_h = \psi(1-m)$, with ψ being the Pareto weight on high-skilled individuals and $(1-m)$ their measure on total population. Correspondingly, we have that $\tilde{\psi}_l = (1-\psi)m$. The Social Planner takes as given the law of motion for preference and technology shocks, given by:

$$Z_t = \rho_z Z_{t-1} + u_{z,t} \quad A_t = \rho_a A_{t-1} + u_{a,t} \quad (\text{A8})$$

$$u_{z,t} = \lambda_z u_{z,t-1} + \sigma_z \varepsilon_{z,t} \quad u_{a,t} = \lambda_a u_{a,t-1} + \sigma_a \varepsilon_{a,t} \quad (\text{A9})$$

$$\begin{bmatrix} \varepsilon_{z,t} \\ \varepsilon_{a,t} \end{bmatrix} \stackrel{\text{i.i.d}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{I}) \quad (\text{A10})$$

The optimality conditions of the planner decision problem are given by:

$$(1-m)c_{h,t} + mc_{l,t} + K_{t+1} - (1-\delta)K_t = K_t^\alpha \left[e^{Z_t} (L_{h,t}^\phi L_{l,t}^{1-\phi}) \right]^{1-\alpha} \quad (\text{A11})$$

$$\frac{\psi(1-m)}{c_{h,t}} = \frac{(1-\psi)m}{c_{l,t}} \quad (\text{A12})$$

$$\frac{1}{c_{j,t}} = \beta \mathbb{E}_t \left[\frac{MPK_{t+1}}{c_{j,t+1}} \right] \quad e^{A_t} b_j h_{j,t}^{\nu^{-1}} = \frac{MPL_{j,t}}{c_{j,t}} \quad \forall j \in \{h, l\} \quad (\text{A13})$$

$$L_{h,t} = (1-m)h_{h,t} \quad L_{l,t} = mh_{l,t} \quad (\text{A14})$$

$$MPK_t = \alpha \frac{Y_t}{K_t} \quad MPL_{h,t} = \phi(1 - \alpha) \frac{Y_t}{L_{h,t}} \quad MPL_{l,t} = (1 - \phi)(1 - \alpha) \frac{Y_t}{L_{l,t}} \quad (\text{A15})$$

We denote by $\theta \in \Theta$ the vector of structural parameters. The optimality conditions in (A8)-(A15) defines a law of motion for the state \mathbf{s}_t and control variables \mathbf{y}_t that we denote by:

$$\mathbf{y}_t = \mathbf{B}(\mathbf{s}_t, \theta) \quad (\text{A16})$$

$$\mathbf{s}_t = \mathbf{\Phi}(\mathbf{s}_{t-1}, \varepsilon_t, \theta). \quad (\text{A17})$$

The function \mathbf{B} and $\mathbf{\Phi}$ are not known in closed form. We approximate them via a first order Taylor expansion around a deterministic steady state of the model (Uribe and Schmitt-Grohe, 2004). Since in the empirical specification of Section 4 we impose a unit root on the preference shock process, we normalize the model's variables to make them stationary. In particular, we scale Y_t , K_{t+1} and $c_{j,t}$ by the factor $e^{-[\phi \frac{v_H}{1+v_H} + (1-\phi) \frac{v_L}{1+v_L}]A_t}$ and hours worked of a j type by the factor $e^{-\frac{v_j}{1+v_j}A_t}$. Denoting by $\tilde{\cdot}$ a normalized variable, one can verify that equations in (A11)-(A15) can be equivalently rewritten as:

$$(1 - m)\tilde{c}_{h,t} + m\tilde{c}_{l,t} + \tilde{K}_{t+1} - (1 - \delta)\tilde{K}_t e^{[\phi \frac{v_H}{1+v_H} + (1-\phi) \frac{v_L}{1+v_L}]A_t} = K_t^\alpha \left[e^{Z_t} (L_{h,t}^\phi L_{l,t}^{1-\phi}) \right]^{1-\alpha} \quad (\text{A18})$$

$$\frac{\psi(1 - m)}{c_{h,t}} = \frac{(1 - \psi)m}{c_{l,t}} \quad (\text{A19})$$

$$\frac{1}{c_{j,t}} = \beta \mathbb{E}_t \left[\frac{MPK_{t+1}}{c_{j,t+1}} \right] \quad e^{A_t} b_j h_{j,t}^{v_j^{-1}} = \frac{MPL_{j,t}}{c_{j,t}} \quad \forall j \in \{h, l\} \quad (\text{A20})$$

$$L_{h,t} = (1 - m)h_{h,t} \quad L_{l,t} = mh_{l,t} \quad (\text{A21})$$

$$MPK_t = \alpha \frac{Y_t}{K_t} \quad MPL_{h,t} = \phi(1 - \alpha) \frac{Y_t}{L_{h,t}} \quad MPL_{l,t} = (1 - \phi)(1 - \alpha) \frac{Y_t}{L_{l,t}} \quad (\text{A22})$$

We use the set of equations (1)-(8) defines a probability distribution for the endogenous variables of our model in terms of a set

We denote by $\mathbf{y}_t = [\log(c_{H,t}), \log(c_{L,t}), \log(h_{H,t}), \log(h_{L,t}), \log(Y_t)]'$ the vector collecting the control variables and by $\mathbf{s}_t = [\log(K_t), A_t, Z_t]'$ the state variables of our model. The above system implicitly defines a probability distribution over $(\mathbf{y}_t, \mathbf{s}_t)$ given by:

$$\mathbf{y}_t = \mathbf{B}(\mathbf{s}_t, \theta), \quad (\text{A23})$$

$$\mathbf{s}_t = \mathbf{\Phi}(\mathbf{s}_{t-1}, \varepsilon_t, \theta). \quad (\text{A24})$$

In the above notation, ε_t is a vector stacking the innovations to preference and technology shocks and θ is the vector of structural parameters. The functions B and Φ are not known in closed form. Our solution method consists in approximating them linearly:

$$\mathbf{y}_t = \mathbf{B}(\theta)\mathbf{s}_t, \tag{A25}$$

$$\mathbf{s}_t = \Phi_{\mathbf{s}}(\theta)\mathbf{s}_{t-1} + \Phi_{\varepsilon}(\theta)\varepsilon_t, \tag{A26}$$

where $\Phi_{\mathbf{s}}(\theta)$, $\Phi_{\varepsilon}(\theta)$ and $\mathbf{B}(\theta)$ are linearization constants derived around a deterministic steady state for our model.¹³ Since the model does not have a unique deterministic steady state when the structural disturbances follow a unit root, some forms of normalization are required. In what follows, we are going to assume that A_t has a unit root while Z_t follows a stationary process. Under these assumptions, the required normalization consists in scaling every model variables i by $e^{\Lambda_{i,a}A_t}$, with $\Lambda_{i,a}$ denoting the long run effect of a preference shock on variable i . This amounts of scaling capital, output and individual consumption by the factor $e^{-\left[\phi\frac{v_H}{1+v_H}+(1-\phi)\frac{v_L}{1+v_L}\right]A_t}$ and hours worked of a j type by the factor $e^{-\frac{v_j}{1+v_j}A_t}$. In fact, performing these operations to the above equilibrium conditions we obtain:

[NORMALIZED OPTIMALITY CONDITIONS HERE]

As we can see, the transformed variables are stationary since the unit root process A_t disappears from the equilibrium conditions.

¹³Quote the paper by Uribe on details for linear approximation of the policy function.

III Proofs and Derivations

III.1 More on VAR representation

Following Fernandez-Villaverde et al. (2006), we assume that the endogenous variables of our model are generated by the system:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{w}_t \quad (\text{A27})$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_{t-1} + \mathbf{D}\mathbf{w}_t \quad (\text{A28})$$

where $\mathbf{y}_t = [\log(L_{1,t}, L_{2,t}, \dots, L_{n-1,t}, L_{n,t})]'$ is the $n \times 1$ vector of labor inputs (control variables), $\mathbf{x}_t = [\log(k_t, Z_{t-1}, A_{t-1})]'$ is the vector of state variables, \mathbf{w}_t is a vector of i.i.d. innovations, and $\mathbf{k}_t = (K_{1,t}, \dots, K_{m,t})$. It is required that \mathbf{y}_t has the same dimension as \mathbf{w}_t .

For the production function:

$$Y_t = Z_t F(K_1, \dots, K_m, L_1, \dots, L_n, \Theta_t) \quad (\text{A29})$$

assume that we observe groups $\tilde{L}_1, \dots, \tilde{L}_l$, which may not coincide with the groups from the production function. Consider the vector $\hat{\mathbf{y}}_t = [\log(F_{\tilde{L}_1,t}, \dots, F_{\tilde{L}_2,t}, \dots, F_{\tilde{L}_l,t}), \mathbf{V}_t]'$, collecting the “marginal products” of the observed groups and other endogenous variables \mathbf{V}_t such that the dimension of $\hat{\mathbf{y}}_t$ equals the dimension of \mathbf{w}_t . Suppose $\tilde{L}_1 = L_1 + L_2$ then the marginal product $F_{\tilde{L}_1,t} = \frac{F_{L_1,t}L_1 + F_{L_2,t}L_2}{L_1 + L_2}$.

Now approximate $\hat{\mathbf{y}}_t$ as a function of the observables and the state variables:

$$\begin{bmatrix} \hat{\mathbf{y}}_t \end{bmatrix} \approx \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} \mathbf{y}_t \\ \mathbf{x}_t \end{bmatrix} \quad (\text{A30})$$

$$= \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} \mathbf{C}\mathbf{x}_{t-1} + \mathbf{D}\mathbf{w}_t \\ \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{w}_t \end{bmatrix} \quad (\text{A31})$$

$$= \underbrace{\begin{bmatrix} X_1\mathbf{C} + X_2\mathbf{A} \end{bmatrix}}_{=: \tilde{\mathbf{C}}} \begin{bmatrix} \mathbf{x}_{t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} X_1\mathbf{D} + X_2\mathbf{B} \end{bmatrix}}_{=: \tilde{\mathbf{D}}} \begin{bmatrix} \mathbf{w}_t \end{bmatrix} \quad (\text{A32})$$

We thus get:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{w}_t \quad (\text{A33})$$

$$\hat{\mathbf{y}}_t = \tilde{\mathbf{C}}\mathbf{x}_{t-1} + \tilde{\mathbf{D}}\mathbf{w}_t \quad (\text{A34})$$

Under the mild assumption that $\tilde{\mathbf{D}}$ is invertible, one has that:

$$\mathbf{w}_t = \tilde{\mathbf{D}}^{-1}(\mathbf{y}_t - \tilde{\mathbf{C}}\mathbf{x}_{t-1}).$$

We can then substitute \mathbf{w}_t in equation (1) and obtain:

$$\left[\mathbf{I} - (\mathbf{A} - \mathbf{B}\tilde{\mathbf{D}}^{-1}\tilde{\mathbf{C}})\mathbf{L} \right] \mathbf{x}_t = \mathbf{B}\tilde{\mathbf{D}}^{-1}\mathbf{y}_t.$$

In the above notation, \mathbf{L} is the lag operator. Under the assumption that the eigenvalues of $(\mathbf{A} - \mathbf{B}\tilde{\mathbf{D}}^{-1}\tilde{\mathbf{C}})$ are less than 1 in modulus, the invertibility condition is satisfied and one can express the vector of unobservable states \mathbf{x}_t through the histories of the observables \mathbf{y}_t . The SVAR literature (long run restrictions, sign restrictions, short run restrictions, etc.) implicitly assume that this invertibility condition holds.

The model, if that condition holds, admits the following VAR(∞) representation for the observables:

$$\hat{\mathbf{y}}_t = \mathbf{C} \left[\mathbf{I} - (\mathbf{A} - \mathbf{B}\tilde{\mathbf{D}}^{-1}\tilde{\mathbf{C}})\mathbf{L} \right]^{-1} \mathbf{B}\tilde{\mathbf{D}}^{-1}\hat{\mathbf{y}}_{t-1} + \tilde{\mathbf{D}}\mathbf{w}_t.$$

It is important to notice that under these assumptions:

$$\mathbb{E} \left[\hat{\mathbf{y}}_t | Z_{t-1}, \{\hat{\mathbf{y}}_j\}_{j=-\infty}^{t-1} \right] = \mathbb{E} \left[\hat{\mathbf{y}}_t | \{\hat{\mathbf{y}}_j\}_{j=-\infty}^{t-1} \right].$$

This implies, among other things, that:

$$\begin{bmatrix} Z_t \\ \hat{\mathbf{y}}_t \end{bmatrix} = \begin{bmatrix} \rho_z & \mathbf{0} \\ \mathbf{0} & \Phi_\infty(L) \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ \hat{\mathbf{y}}_{t-1} \end{bmatrix} + \begin{bmatrix} \sigma_z \varepsilon_{z,t} \\ \tilde{\mathbf{D}}\mathbf{w}_t \end{bmatrix}.$$

This justifies in the limit ($p \rightarrow \infty$) Assumption 1 and Assumption 3 in the main text.

III.2 Proof of Lemma 1

Suppose that the dimension of \mathbf{Y}_t is $n - k - 1$. In order to check that our state space is minimal, one need to verify the observability and controllability conditions are satisfied in our state space model. The observability matrix is given by:

$$\mathcal{O}_{(n-1)^2 \times n}^n = \begin{bmatrix} \mathbf{B}_{(n-1) \times n} \\ \mathbf{B}\Phi^{\circ}_{(n-1) \times n} \\ \dots \\ \mathbf{B}\Phi^{\circ, n-1}_{(n-1) \times n} \end{bmatrix}.$$

The observability condition is satisfied if $\mathcal{O}_{(n-1)^2 \times n}^n$ is of full rank. First notice that \mathbf{B} is of rank $n - 1$. Now, suppose that the observability condition is violated. That would imply that one can find a set of coefficients, that we collect in the matrix \mathbf{A} with the property that:

$$\mathbf{A}\mathbf{B} = \mathbf{B}\Phi^{\circ}.$$

Given our knowledge of the \mathbf{B} matrix, that would imply the following relation between \mathbf{A} and Φ° :

$$\sum_{l=1}^k a_{j,l} = \rho_z \quad \forall j < k + 1 \quad \sum_{l=1}^k a_{j,l} = \rho_z \quad \forall j \geq k$$

$$\mathbf{A} = \Phi^{\circ}$$

Notice that Assumption 2 in the main text guarantees that $\sum_{l=1}^k a_{j,l} \neq \rho_z$ for at least one $j < k + 1$. Thus, by contradiction we must have that $\mathbf{B}\Phi^{\circ}$ can not be written as a linear combination of \mathbf{B} . The observability matrix is of full rank. The controllability matrix in our state space system is of full rank as long as the variance covariance error of the state variables is positive definite.

As a result, our state space realization is observable and controllable, hence minimal.

III.3 Proof of Proposition 1

Without loss of generality, assume that \mathbf{S}_t follows a VAR(1). From the restriction:

$$\mathbf{B}\mathbf{T} = \mathbf{B}$$

one gets that the set of admissible \mathbf{T} matrices satisfies:

$$\begin{bmatrix} \tilde{\mathbf{T}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 1 + \kappa_1 & \kappa_2 & \dots & \kappa_i & \dots & \kappa_{k+1} & \kappa_{k+2} & \kappa_{k+3} & \dots & \kappa_n & \kappa_{n+1} \\ \kappa_1 & 1 + \kappa_2 & \dots & \kappa_i & \dots & \kappa_{k+1} & \kappa_{k+2} & \kappa_{k+3} & \dots & \kappa_n & \kappa_{n+1} \\ \dots & \dots \\ \kappa_1 & \kappa_2 & \dots & \kappa_i & \dots & \kappa_{k+1} & \kappa_{k+2} & \kappa_{k+3} & \dots & \kappa_n & \kappa_{n+1} \\ \kappa_1 & \kappa_2 & \dots & \kappa_i & \dots & 1 + \kappa_{k+1} & \kappa_{k+2} & \kappa_{k+3} & \dots & \kappa_n & \kappa_{n+1} \\ 0 & 0 & \dots & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix},$$

where m is the dimension of \mathbf{Y}_t .

The matrix \mathbf{T} has to satisfy also the system of equations:

$$\hat{\Phi} = \mathbf{T}\Phi^0\mathbf{T}^{-1}, \quad (\text{A35})$$

Given the above characterization for \mathbf{T} , one has:

$$\mathbf{T}^{-1} = \begin{bmatrix} 1 - \frac{\kappa_1}{\kappa} & -\frac{\kappa_2}{\kappa} & \dots & -\frac{\kappa_i}{\kappa} & \dots & -\frac{\kappa_{k+1}}{\kappa} & \frac{-\kappa_{k+2}}{\kappa} & \frac{-\kappa_{k+3}}{\kappa} & \dots & \frac{-\kappa_n}{\kappa} & \frac{-\kappa_{n+1}}{\kappa} \\ -\frac{\kappa_1}{\kappa} & 1 - \frac{\kappa_2}{\kappa} & \dots & -\frac{\kappa_i}{\kappa} & \dots & -\frac{\kappa_{k+1}}{\kappa} & \frac{-\kappa_{k+2}}{\kappa} & \frac{-\kappa_{k+3}}{\kappa} & \dots & \frac{-\kappa_n}{\kappa} & \frac{-\kappa_{n+1}}{\kappa} \\ \dots & \dots \\ -\frac{\kappa_1}{\kappa} & -\frac{\kappa_2}{\kappa} & \dots & -\frac{\kappa_i}{\kappa} & \dots & -\frac{\kappa_{n+1}}{\kappa} & \frac{-\kappa_{k+2}}{\kappa} & \frac{-\kappa_{k+3}}{\kappa} & \dots & \frac{-\kappa_n}{\kappa} & \frac{-\kappa_{n+1}}{\kappa} \\ -\frac{\kappa_1}{\kappa} & -\frac{\kappa_2}{\kappa} & \dots & -\frac{\kappa_i}{\kappa} & \dots & 1 - \frac{\kappa_{n+1}}{\kappa} & \frac{-\kappa_{k+2}}{\kappa} & \frac{-\kappa_{k+3}}{\kappa} & \dots & \frac{-\kappa_n}{\kappa} & \frac{-\kappa_{n+1}}{\kappa} \\ 0 & 0 & \dots & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix},$$

where $\kappa = 1 + \kappa_1 + \kappa_2 + \dots + \kappa_{k+1}$. Let $\lambda_{ij} = (\mathbf{\Phi}^o \mathbf{T}^{-1})_{ij}$. Then:

$$\lambda_{ij} = \phi_{ij}^o + \sum_{l=1}^{k+1} \phi_{il}^o \frac{-\kappa_l}{\kappa} \quad (\text{A36})$$

and thus

$$\gamma_{ij} = \lambda_{ij} + \sum_{l=1}^{n+1} \kappa_l \lambda_{lj}, \text{ if } i \leq n+1 \quad (\text{A37})$$

$$\gamma_{ij} = \lambda_{ij}, \text{ if } i > n+1. \quad (\text{A38})$$

We are interested in the first row and column of this matrix, γ_{1j} and γ_{j1} . Assumption 3 implies that:

$$\gamma_{1j} = 0, \text{ for } j > 1 \quad (\text{A39})$$

$$\gamma_{j1} = 0, \text{ for } j > 1. \quad (\text{A40})$$

We have that

$$\lambda_{j1} = \sum_{l=1}^{k+1} \phi_{jl}^o \frac{-\kappa_l}{\kappa} \quad (\text{A41})$$

and using restrictions on $\hat{\mathbf{\Phi}}$:

$$\lambda_{j1} = \sum_{l=1}^{k+1} \phi_{jl}^o \frac{-\kappa_l}{\kappa} \text{ if } j \geq 2 \quad (\text{A42})$$

$$\lambda_{11} = \phi_{11}^o + \sum_{l=1}^{k+1} \phi_{1l}^o \frac{-\kappa_l}{\kappa} \quad (\text{A43})$$

$$= \phi_{11}^o - \frac{\kappa_1}{\kappa} \phi_{11}^o \text{ if } j = 1. \quad (\text{A44})$$

Therefore:

$$\gamma_{11} = \lambda_{11} + \sum_{l=1}^{n+1} \kappa_l \lambda_{l1} \quad (\text{A45})$$

$$= \phi_{11}^o - \frac{\kappa_1}{\kappa} \phi_{11}^o + \phi_{11}^o \kappa_1 + \sum_{l=1}^{n+1} \kappa_l \sum_{m=1}^{k+1} \phi_{lm}^o \frac{-\kappa_m}{\kappa} \quad (\text{A46})$$

$$\gamma_{j1} = \lambda_{j1} + \sum_{l=1}^{n+1} \kappa_l \lambda_{l1} \quad (\text{A47})$$

$$= \sum_{l=1}^{k+1} \phi_{jl}^o \frac{-\kappa_l}{\kappa} + \phi_{11}^o \kappa_1 + \sum_{l=1}^{n+1} \kappa_l \sum_{m=1}^{k+1} \phi_{lm}^o \frac{-\kappa_m}{\kappa}, \text{ if } 2 \leq j \leq k+1 \quad (\text{A48})$$

$$\gamma_{j1} = \lambda_{j1} = \sum_{l=1}^{k+1} \phi_{jl}^o \frac{-\kappa_l}{\kappa}, \text{ if } j > k+1. \quad (\text{A49})$$

Considering the first row, we have that:

$$\lambda_{1j} = \phi_{1j}^o + \sum_{l=1}^{k+1} \phi_{1l}^o \frac{-\kappa_l}{\kappa} \quad (\text{A50})$$

$$= -\phi_{11}^o \frac{\kappa_j}{\kappa} \quad \text{if } j \geq 2 \quad (\text{A51})$$

$$\lambda_{11} = \phi_{11} + \sum_{l=1}^{k+1} \phi_{1l} \frac{-\kappa_j}{\kappa} \quad (\text{A52})$$

$$= \phi_{11}^o - \phi_{11}^o \frac{\kappa_1}{\kappa} \quad (\text{A53})$$

and thus:

$$\gamma_{11} = \lambda_{11} + \sum_{l=1}^{n+1} \kappa_l \lambda_{l1} \quad (\text{A54})$$

$$= \phi_{11}^o - \phi_{11}^o \frac{\kappa_1}{\kappa} + \phi_{11}^o \kappa_1 + \sum_{l=1}^{n+1} \kappa_l \sum_{m=1}^{k+1} \phi_{lm}^o \frac{-\kappa_1}{\kappa} \quad \text{if } 2 \leq j \leq n+1 \quad (\text{A55})$$

$$\gamma_{1j} = \lambda_{1j} + \sum_{l=1}^{n+1} \kappa_l \lambda_{lj} \quad (\text{A56})$$

$$= -\phi_{11}^o \frac{\kappa_j}{\kappa} + \sum_{l=1}^{n+1} \kappa_l \left(\phi_{lj}^o + \sum_{m=1}^{k+1} \phi_{lm}^o \frac{-\kappa_j}{\kappa} \right) \quad \text{if } j \leq n+1 \quad (\text{A57})$$

$$\gamma_{1j} = \lambda_{1j} = -\phi_{11}^o \frac{\kappa_j}{\kappa} \quad \text{if } j > n+1. \quad (\text{A58})$$

Equation (A58) implies that $\kappa_j = 0$ if $j > k+1$.

If $m \geq 1$ (that is we have endogenous variables) and for some $j > k+1$, $\sum_{l=1}^{k+1} \phi_{jl}^o \neq 0$ then equation (A49) implies that $\kappa_1 = 0$. If $m = 0$ then we have to use equation (A48) to be able to say something about κ_1 . Suppose $\kappa_1 \neq 0$ then equation (A48) implies that $\sum_{l=1}^{k+1} \phi_{jl}^o$ is constant for all $2 \leq j \leq k+1$. If this is not the case, $\kappa_1 = 0$. Assumption 1 rules out this last scenario, so $\kappa_1 = 0$.

What remains to be considered are $\kappa_2, \dots, \kappa_{k+1}$. To this aim, we use equation (A57) and that $\kappa_j = 0$ if $j > k+1$ so that for each $2 \leq j \leq k+1$

$$0 = -\phi_{11}^o \kappa_j + \sum_{l=1}^{k+1} \kappa_l \kappa \phi_{lj}^o - \kappa_j \sum_{l=1}^{k+1} \kappa_l \sum_{m=1}^{k+1} \phi_{lm}^o \quad (\text{A59})$$

$$= -\phi_{11}^o \kappa_j + (1 + \kappa_1 + \kappa_2 + \dots + \kappa_{k+1}) \sum_{l=1}^{k+1} \kappa_l \phi_{lj}^o - \kappa_j \sum_{l=1}^{k+1} \kappa_l \sum_{m=1}^{k+1} \phi_{lm}^o \quad (\text{A60})$$

Using also that $\kappa_1 = 0$, the gradient w.r.t. $\kappa_2 = 0, \dots, \kappa_{n+1} = 0$ equals

$$(\phi_{2j}, \phi_{3j}, \dots, \phi_{jj} - \phi_{11}, \dots, \phi_{k+1j}). \quad (\text{A61})$$

Thus the derivative(-matrix) equals

$$\begin{bmatrix} \phi_{22}^o - \phi_{11}^o & \phi_{32}^o & \dots & \phi_{j2}^o & \dots & \phi_{n2}^o & \phi_{n+12}^o \\ \phi_{23}^o & \phi_{33}^o - \phi_{11}^o & \dots & \phi_{j3}^o & \dots & \phi_{n3}^o & \phi_{n+13}^o \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \phi_{2n}^o & \phi_{3n}^o & \dots & \phi_{jn}^o & \dots & \phi_{n3}^o - \phi_{11}^o & \phi_{n+13}^o \\ \phi_{2n+1}^o & \phi_{3n+1}^o & \dots & \phi_{jn+1}^o & \dots & \phi_{nn+1}^o & \phi_{n+13}^o - \phi_{11}^o \end{bmatrix},$$

If the Φ^o matrix has full rank and ϕ_{11}^o is not an eigenvalue then this matrix has full rank too. This implies that the set of equations is locally invertible what is equivalent to being locally identified.

We finally show that the assumption of VAR(1) is without loss of generality. First of all, let's define $\Gamma = [-\mathbf{1}, \mathbf{I}_k]$ and let's rewrite the transition equation as:

$$\mathbf{S}_t = \Phi \mathbf{S}_{t-1} + \begin{bmatrix} \Sigma^{1/2} \\ \mathbf{0} \end{bmatrix} \mathbf{e}_t$$

where $\mathbf{S}_t = [\mathbf{S}_t^u, \mathbf{S}_t^o, \mathbf{S}_{t-1}^u, \mathbf{S}_{t-1}^o, \dots, \mathbf{S}_{t-p+1}^u, \mathbf{S}_{t-p+1}^o]'$ (u stands for unobserved, o stands for observed). First of all, notice that since:

$$\begin{bmatrix} \hat{\Sigma}^{1/2} \\ \mathbf{0} \end{bmatrix} = \mathbf{T}_p \begin{bmatrix} \Sigma^{1/2} \\ \mathbf{0} \end{bmatrix}$$

The \mathbf{T}_p matrix must have its lower-left block equal to $\mathbf{0}$ ($\mathbf{S}_t^u, \mathbf{S}_t^o$ can not depend on $\mathbf{S}_{t+k}^u, \mathbf{S}_{t+k}^o$ for $k > 0$).

Also, by using the restrictions on the \mathbf{T}_p matrix implied by the measurement equation,

we obtain:

$$\mathbf{T}_p = \begin{bmatrix} \mathbf{I} + \mathbf{K}_{11}^u & \mathbf{K}_{11}^o & \mathbf{K}_{12}^u & \mathbf{K}_{12}^o & \dots & \mathbf{K}_{1p}^u & \mathbf{K}_{1p}^o \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} + \mathbf{K}_{22}^u & \mathbf{K}_{22}^o & \dots & \mathbf{K}_{2p}^u & \mathbf{K}_{2p}^o \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \dots & \mathbf{0} & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} + \mathbf{K}_{pp}^u & \mathbf{K}_{pp}^o \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{I} \end{bmatrix}$$

and where the above matrices solve the set of equations:

$$\mathbf{\Gamma} \mathbf{K}_{ij}^u = \mathbf{\Gamma} \mathbf{K}_{ij}^o = \mathbf{0}.$$

Now, notice that from the above \mathbf{T}_p matrix we obtain for $k = 1, \dots, p$:

$$\hat{\mathbf{S}}_{t-k+1}^u = \mathbf{S}_{t-k+1}^u + \sum_{m=k}^p \mathbf{K}_{km}^u \mathbf{S}_{t-m+1}^u + \mathbf{K}_{km}^o \mathbf{S}_{t-m+1}^u.$$

Using the equation for $\hat{\mathbf{S}}_t^u$ at different lags yields:

$$\hat{\mathbf{S}}_t^u = \mathbf{S}_t^u + \sum_{m=1}^p \mathbf{K}_{1m}^u \mathbf{S}_{t-m+1}^u + \mathbf{K}_{1m}^o \mathbf{S}_{t-m+1}^u \quad (\text{A62})$$

$$= \mathbf{S}_t^u + \sum_{m=2}^p \mathbf{K}_{2m}^u \mathbf{S}_{t-m+2}^u + \mathbf{K}_{2m}^o \mathbf{S}_{t-m+2}^u \quad (\text{A63})$$

$$= \dots \quad (\text{A64})$$

$$= \mathbf{S}_t^u + \sum_{m=p}^p \mathbf{K}_{pm}^u \mathbf{S}_{t-m+p}^u + \mathbf{K}_{pm}^o \mathbf{S}_{t-m+p}^u. \quad (\text{A65})$$

Taking the difference between equation $k \leq p-1$ and equation p yields

$$0 = \sum_{m=k}^p \mathbf{K}_{km}^u \mathbf{S}_{t-m+k}^u + \mathbf{K}_{km}^o \mathbf{S}_{t-m+k}^u - (\mathbf{K}_{pp}^u \mathbf{S}_t^u + \mathbf{K}_{pp}^o \mathbf{S}_t^u) \quad (\text{A66})$$

$$= (\mathbf{K}_{kk}^u - \mathbf{K}_{pp}^u) \mathbf{S}_t^u + (\mathbf{K}_{kk}^o - \mathbf{K}_{pp}^o) \mathbf{S}_t^u + \sum_{m=k+1}^p \mathbf{K}_{km}^u \mathbf{S}_{t-m+k}^u + \mathbf{K}_{km}^o \mathbf{S}_{t-m+k}^u. \quad (\text{A67})$$

Controllability of our state space (see Lemma 1) guarantees that for every vector \mathbf{x} (of the same dimension as the state vector) we have a sequence of shocks such that

$\mathbf{x} = [\mathbf{S}_t^u, \mathbf{S}_t^o, \mathbf{S}_{t-1}^u, \mathbf{S}_{t-1}^o, \dots, \mathbf{S}_{t-p+1}^u, \mathbf{S}_{t-p+1}^o]'$. In particular, this is true for the vector $\mathbf{x} = [0, \dots, 1, \dots, 0]'$. Using these insights in the above equation, we see that for any element κ of any of these matrices we get the equation $0 = \kappa$. We thus get that

$$\mathbf{K}_{kk}^u = \mathbf{K}_{pp}^u \text{ for all } k \quad (\text{A68})$$

$$\mathbf{K}_{kk}^o = \mathbf{K}_{pp}^o \text{ for all } k \quad (\text{A69})$$

$$\mathbf{K}_{km}^u = \mathbf{K}_{km}^o = 0 \text{ for all } k \text{ and } m > k \quad (\text{A70})$$

This means for the transformation matrix \mathbf{T}_p :

$$\mathbf{T}_p = \begin{bmatrix} \mathbf{I} + \mathbf{K}_{11}^u & \mathbf{K}_{11}^o & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} + \mathbf{K}_{22}^u & \mathbf{K}_{22}^o & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \dots & \mathbf{0} & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} + \mathbf{K}_{pp}^u & \mathbf{K}_{pp}^o \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{I} \end{bmatrix}$$

or equivalently

$$\mathbf{T}_p = \begin{bmatrix} \mathbf{T}_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_1 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{T}_1 & \dots & \mathbf{0} & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{T}_1 \end{bmatrix}$$

where \mathbf{T}_1 is the transformation matrix for the case with one lag.

As is the VAR(1) case we know that the off-diagonal elements of the first row and the first column of $\hat{\Phi}$ and $\mathbf{T}_1 \Phi^o \mathbf{T}_1^{-1}$ are zero and thus the same arguments as above apply. We therefore have that \mathbf{T}_1 is the identity matrix and thus is \mathbf{T}_p .