

Federal Reserve Bank of Minneapolis
Research Department Staff Report 315

Revised November 2008

Does Neoclassical Theory Account for the Effects of Big Fiscal Shocks? Evidence from World War II*

Ellen R. McGrattan

Federal Reserve Bank of Minneapolis
and University of Minnesota

Lee E. Ohanian

University of California, Los Angeles
and Federal Reserve Bank of Minneapolis

ABSTRACT

There is much debate about the usefulness of the neoclassical growth model for assessing the macroeconomic impact of fiscal shocks. We test the theory using data from World War II, which is by far the largest fiscal shock in the history of the United States. We take observed changes in fiscal policy during the war as inputs into a parameterized, dynamic general equilibrium model and compare the values of all variables in the model to the actual values of these variables in the data. Our main finding is that the theory quantitatively accounts for macroeconomic activity during this big fiscal shock.

*We thank the National Science Foundation for financial support. We also thank Casey Mulligan, Cristina DeNardi, Ed Prescott, Julio Rotemberg, and three referees for helpful comments on earlier drafts. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Introduction

World War II is the largest fiscal shock in the history of the United States, and it also represents the most significant economic boom in U.S. history. Many economists agree that wartime government spending contributed to this economic boom. But beyond this general level of agreement, there is significant debate about the impact of World War II—and more generically the impact of other large fiscal changes—on the economy. A major point of disagreement is the appropriate theoretical framework for quantitatively studying the impact of World War II on the economy. Some economists argue that the standard neoclassical growth model is a useful tool for accounting for World War II, while others argue that accounting for the impact of such large fiscal changes requires major departures from the neoclassical framework.¹

The disagreement about the usefulness of the neoclassical model for understanding World War II partially reflects the fact that there is no comprehensive analysis of the impact of the World War II shock on the U.S. economy using the neoclassical model. Consequently, there are several open questions: What are the neoclassical model's successes and failures in accounting for the World War II economy? What are the impacts of the other large World War II changes, such as the very large changes in tax rates, in government investment, and in the draft? World War II is the biggest macroeconomic shock to hit the U.S. economy and therefore provides a unique opportunity to test the neoclassical model.

We address these questions using a neoclassical growth model tailored to study World War II by including four types of exogenous variables that were important during the war: (1) government spending, (2) income taxes, (3) the draft, and (4) productivity shocks. We conduct a sequence of quantitative experiments that investigate how well the model accounts for the major macroeconomic variables: output and its components, hours worked, and factor returns during World War II.

We find that real GNP, investment, consumption, labor supply, and the returns to capital and labor from the model are similar to those in the data; the model captures the large increase in real output and hours worked, the large decline in investment, and the wartime changes in factor prices. Regarding the relative importance of the different wartime factors, we find that the most important change by far is the large increase in government spending, followed by productivity. The very large changes in the draft, and in capital and labor income tax rates have much smaller effects.

The paper is organized as follows. Section 2 reviews the literature and defines the scope of

our analysis. Section 3 summarizes the changes in the exogenous variables that we feed into the model. Section 4 summarizes the economic restrictions and regulations adopted during World War II. Section 5 presents the model and model parameterization. Section 6 conducts the quantitative analysis by comparing the model to the data. Section 7 discusses our findings in light of the issues raised in the literature. Section 8 summarizes the sensitivity analysis we conduct. Section 9 summarizes and concludes.

2. Literature Review and the Scope of Analysis

We are unaware of any comprehensive assessment of the impact of World War II on the U.S. economy, particularly any studies that systematically address all of the open questions about the neoclassical model's ability to account for the wartime economy. Specifically, there are questions about the conformity of all of the key variables in the growth model: the World War II boom in output and labor, the components of output (consumption and investment), and changes in pre- and post-tax factor prices and returns.² These are all open questions because none of the studies cited below simulate the model's response to the World War II shock to compare the model's equilibrium path to U.S. data. This paper will address these questions.

There are a number of questions about labor supply and after-tax wages. Mulligan (1998) and Baxter and King (1993) question the ability of the neoclassical model to account for labor supply and after-tax factor price changes during World War II. They are motivated by the fact that two opposing forces impact hours worked during wars. Higher government spending will tend to increase labor, but higher taxes may decrease labor. Mulligan focuses his analysis on World War II. He conjectures that the neoclassical model cannot plausibly account for the large increase in World War II labor supply, as he shows that after-tax wages and returns to capital during the war are not particularly high. Mulligan concludes that the model would require patriotism, modeled as preference shifts, to jointly account for wartime labor supply and factor prices. Baxter and King raise a similar concern, questioning whether the model can account for the boom in labor supply given the large tax increases that occurred during World War II. Burnside, Eichenbaum, and Fisher (2004) question the ability of the standard neoclassical model to account for macroeconomic changes during post-World War II military episodes, rather than World War II. They argue that the model needs to be augmented with habit formation and investment adjustment costs.

There are also questions about pre-tax wages. Rotemberg and Woodford (1992) focus on the impact of military spending shocks on pre-tax real wages. They note that the neoclassical

model, with constant returns to scale production and perfectly competitive product and factor markets, can only account for higher pre-tax real wages during World War II with a shift in the production function, through either capital accumulation and/or a technological shift. This leads them to conjecture that the model requires significant, time-varying markups to account for pre-tax wages.³ We will therefore compare pre-tax real wages from our competitive model to actual pre-tax real wages to address this issue.

Finally, there are questions about changes in the composition of output between consumption and investment during periods of large fiscal shocks. Blanchard and Perotti (2002) use a vector autoregression (VAR) to analyze the impact of post-World War II military spending shocks. They do not focus their analysis on any particular theoretical model, but argue that their VAR results regarding the decrease in private investment in response to military spending shocks are a challenge for most theories. We will therefore compare the division of output between investment and consumption to address questions about the response of the components of output to the war shock. Another reason to examine the division of output is because the composition of expenditure changed significantly during the war. Employment in motor vehicles and housing (consumer goods) fell substantially, while employment in other sectors, such as steel and chemicals, expanded considerably. These compositional changes at the industry level may potentially shift the distribution of output at the aggregate level between consumption and investment.

All of these questions fall under the umbrella of the broad issue that we address here: what are the quantitative successes and failures of the neoclassical model for understanding the World War II macroeconomy? We address this question (and by implication, the other questions cited above) as follows. We first identify the time series of the wartime changes in exogenous variables, feed these variables into a parameterized model, compute the equilibrium, and then graphically compare the actual time series of each endogenous variable (output and its components, labor, and factor prices) to their model counterparts for the years 1939–1946.

We focus our analysis on World War II because it is the most important fiscal shock in the history of the United States; the shocks are by far the largest and are also very likely to be exogenous. These features of World War II provide a unique environment for testing the neoclassical model. Other military episodes, such as the Vietnam War or the defense buildup during the Carter and Reagan administrations, are much smaller in terms of changes in spending, taxes, and the draft. These other episodes are certainly interesting to study but are beyond the scope of this paper.

3. Exogenous Variables

This section summarizes the exogenous variables that we include in the model. The three types of government policy (fiscal) variables are : government spending on consumption and investment goods, income tax rates, and the draft. We call this the *fiscal state* of the economy. In addition to these fiscal variables, we also include productivity shocks. Figure 1 displays all six exogenous variables. The most striking feature of these data is their magnitude. These large changes provide an excellent testing ground for evaluating the neoclassical model.

The first series in Figure 1 is our measure of real government consumption and is total government spending less government investment and military compensation, all real. The source of the government spending data is the U.S. Department of Commerce (1975, Series F167; 1986, Tables 1.2, 3.7A, and 3.8A; and 1987, Table B12). The second series in Figure 1 is our measure of real government investment in plant and equipment that is used in the production of goods and services. This investment is total investment by federal, state, and local government less investment in military equipment and structures. Much of the government investment in this period was in government-owned, privately operated capital. This capital investment in equipment and factories substituted for private investment. We divide the series for government consumption and investment by the population over 16 from U.S. Department of Commerce (1975, Series A39) and by the growth rate of labor-augmenting technology, which we estimate at 2 percent per year. To put the expenditure series in interpretable units, we also divide them by nonmilitary output in 1946, where nonmilitary output is gross national product (GNP) less military compensation.

The figure shows that both categories of government spending rose significantly over the course of the war. At the beginning and end of the war, government consumption is about 20 percent of trend nonmilitary output. At its peak, government consumption rises to almost 50 percent of trend nonmilitary output. Nonmilitary government investment doubles between 1941 and 1942, reaching almost 9 percent of trend nonmilitary output.

The third and fourth series in Figure 1 are estimates of labor and capital income tax rates from Joines (1981, Series MTRL1 and MTRK1). The most striking feature is that tax rates rose significantly. Labor taxes more than doubled between 1939 and 1945, rising from about 8 percent to about 20 percent. Capital tax rates rose about 50 percent, from a little over 40 percent to more than 60 percent. The fifth series in Figure 1 shows the fraction of the working-age population in the military. The fraction rises from less than 2 percent before the war to more than 11 percent at the

peak of the war.⁴ These data clearly show that all of these elements of fiscal policy—government spending, labor income tax rates, capital income tax rates, and the number of individuals drafted—rose substantially during the war.

In addition to these fiscal variables, we also include productivity shocks. The last series in Figure 1 shows detrended total factor productivity (TFP). We measure TFP using a capital share of 0.38 (we include both government and private capital) and a labor share of 0.62. Output used in the calculation of TFP is GNP less military compensation, which we refer to as *nonmilitary output*. The source of the GNP data is the U.S. Department of Commerce (1986, Table 1.2). The capital stock is the sum of private and public capital used in producing nonmilitary output; we exclude military equipment and structures. The source of the capital stock data is U.S. Department of Commerce (1987, Tables A6, A9, A15, A17, and A19). The labor input in the calculation of TFP is nonmilitary manhours reported in Kendrick (1961, Table A-X). We detrend TFP at its average growth rate of 1.2 percent per year. Detrended TFP rises 13 percent between 1941 and 1945 and 25 percent between 1939 and 1945. After 1945, TFP falls.

There are a number of reasons why the increase in TFP is indeed plausible. The literature notes several factors, including significant increases in research and development spending (R&D), the development of management science procedures and operations research practices, and substantial government infrastructure investment, which significantly raised TFP above trend during this period.

Regarding R&D spending, Mowery and Rosenberg (2000) document that real federal R&D expenditures rose from \$83 million in 1930 to \$1.31 billion (in 1930 dollars) in 1945.⁵ This large increase in spending, which was concentrated among leading firms in science-based industries and research universities, plausibly led to significant productivity advances in a variety of industries. R&D grants were primarily managed by the Office of Scientific Research and Development, which entered into research partnerships with many leading universities and corporations, including 75 contracts with MIT. In conjunction with these grants, members of the scientific community were mobilized to recommend, guide, and participate in scientific research. Mowery and Rosenberg note that this advent of federal funding of R&D was the precursor of postwar federally funded and subsidized R&D programs.

Mowery and Rosenberg (2000) note that these R&D expenditures raised productivity in a number of manufacturing industries, including airframes, shipping, radar, microwave technology,

and fertilizer. Similarly, Davies and Stammers (1975) report that other industries with significant advances included air travel, synthetic rubber, oxygen steel, titanium, jet propulsion, silicones, urethanes, polythenes, chemotherapy, polymers, insecticides, nylon, and Teflon. Bullard (1975) reports large advances in electronics and instrumentation.

Kendrick (1961) reports output per hour and TFP for several of these industries for selected years, including 1937, 1948, and 1953. Regarding the 1937–1948 period, TFP in the manufacturing sector rose 19 percent (with no adjustment made for trend growth). The gains are relatively high in several of the sectors that were affected by these R&D programs, including chemicals and allied products (up 49 percent), electrical machinery (up 26 percent), miscellaneous and instruments (up 25 percent), and primary metals (up 42 percent). These gains between 1937 and 1948, however, probably represent a lower bound on an estimate of technological change during the war, because the postwar conversion of the economy likely temporarily lowered productivity. For example, Kendrick (1961) reports that economy-wide TFP fell by 2.3 percent between 1946 and 1948, and rose by 13 percent between 1948 and 1953. Moreover, it is likely that the productivity decline between 1946 and 1948 was larger in the manufacturing sector, because that sector was relatively more affected by the reconversion. This suggests calculating productivity change between 1937 and 1953, which is the next available year after 1948, as an alternative measure of wartime technological change. This is probably an upper bound, given that it is seven years after World War II. Between 1937 and 1953, manufacturing TFP is about 34 percent higher, and TFP in the other sectors cited above are up about 83 percent, 61 percent, 45 percent, and 45 percent, respectively. Calculations made over both of these time intervals indicate that productivity growth in these four sectors that were particularly impacted by R&D activity is well above average.

The literature also reports other sources of productivity advances. Davies and Stammers (1975) discuss that a significant source of World War II productivity growth was the development of management science and operations research practices by industrial scientists. These practices led to increased efficiency in factory output and more broadly in organizations. Field (2003) cites other factors that raised productivity during the war, including significant government infrastructure investment in roads, highways, bridges, and airports during this period. Moreover, he notes that the very high levels of private R&D spending of the 1930s likely continued to have productivity spillovers into the 1940s. Finally, Alchian (1963), among several other economists, has argued that wartime learning-by-doing raised productivity considerably.

As Figure 1 shows, total factor productivity declined after the war was over. Gordon (1969),

and in his discussion of Braun and McGrattan (1993), states that this decline is empirically plausible because a substantial fraction of plant and equipment that was built during the war for military purposes was converted to private production after the war. This conversion effort, which included the disassembling of some equipment and plants, as well as the conversion of plants to produce different goods, likely reduced productivity, particularly in the manufacturing sector, in the immediate postwar period. The fact that productivity rose sharply in 1949 and 1950 is also consistent with this view.

While quantitatively accounting for the contribution of all of these factors to aggregate productivity change both during and after the war is beyond the scope of this study, this evidence does indicate that total factor productivity growth was higher than normal during the war and that productivity declined temporarily after the war. Our baseline approach will treat this productivity change as exogenous and assess its implications for the wartime economy. Since changes in capacity utilization could account for some of the change in the Solow residual during this period, we will consider a version of our model with variable capacity in our sensitivity analyses.

4. Wartime Economic Restrictions and Regulations

A number of economic restrictions and regulations were adopted during World War II that are not present during peacetime. Major restrictions include nominal wage and price controls and rationing (through ration coupons) of meat, butter, gasoline, sugar and other nondurable goods, de facto rationing of some durable goods, such as autos and residential capital, which were produced in very small quantities during the war, and Federal Reserve management and control of U.S. Treasury security markets that fixed the nominal interest rates on these securities between $3/8$ of 1 percent for short-term Treasury debt to 2.5 percent for long-term Treasury debt.⁶

Whether these restrictions had quantitatively important effects on the major macroeconomic aggregates during this period is an open question because there is no work addressing this question within the growth model. The impact of these restrictions is also uncertain because economic agents found ways to get around at least some of these restrictions. For example, the trading of ration coupons and the emergence of black markets allowed households to at least partially offset the impact of rationing, and firms supplemented salaries with non-wage benefits to offset the impact of wage controls (see Rockoff 1984).

A model that captured these regulations and restrictions would differ considerably from the standard growth model we are interested in testing. For example, understanding the apparent

success of the Federal Reserve’s interest rate fixing policy would require a model that included the many restrictions on the types of assets that could be held by regulated intermediaries, because almost all federal debt was held by these institutions.⁷ Including these regulations would clearly add significantly to the complexity of the model and would preclude us from assessing how well the standard model does. We therefore abstract from these financial regulations and restrictions.

We also abstract from monetary factors because monetary policy was fundamentally tied to the Treasury’s financing of the war through the pegging of interest rates. We therefore leave this very interesting and challenging issue for further research.

Here we assess how well a simple model without coupon rationing and price and wage controls, without constrained investment in certain sub-sectors of the economy, and without control of the pricing of government debt/monetary policy can account for the U.S. data during World War II. This simple approach is a natural first step in assessing World War II, because if any of these abstractions are quantitatively important, then the model will fail noticeably along one or more dimensions.

5. Model Economy

We start with a standard neoclassical model and tailor it to study the impact of wartime shocks. The model includes government consumption, government investment in physical capital, government payments to military personnel, taxation of capital and labor income, and the draft.

There is an infinitely lived representative family with two types of family members: “civilians” and “draftees.” Both types of family members have identical preferences given by $U(c, l) = \log(c) + V(1 - l)$, where V is a concave and continuously differentiable function. There are N_t total family members in period t , with fraction a_t who are in the military and fraction $(1 - a_t)$ who are civilians. This specification allows us to distinguish between civilian workers available for production $(1 - a_t)$, and the hours worked by the civilians (l_{ct}).

In each period t , the family optimally chooses consumption of both types, which we denote by c_{ct} and c_{dt} , respectively. The family also chooses private investment in physical capital, i_{pt} , civilian labor input, l_{ct} , and next period government bond holdings b_{t+1} to maximize its lifetime utility. Labor input for family members in the military is exogenously fixed at \bar{l}_d . The family’s maximization problem is given by

$$(1) \quad \max_{\{c_{ct}, c_{dt}, i_{pt}, l_{ct}, b_{t+1}\}} \sum_{t=0}^{\infty} \beta^t \{(1 - a_t)U(c_{ct}, l_{ct}) + a_t U(c_{dt}, \bar{l}_d)\} N_t$$

subject to

$$\begin{aligned}
(2) \quad & (1 - a_t)c_{ct} + a_t c_{dt} + i_{pt} + b_{t+1} \\
& = (1 - \tau_{kt})r_{pt}k_{pt} + (1 - \tau_{lt})w_t(1 - a_t)l_{ct} + \tau_{kt}\delta k_{pt} + R_t b_t + (1 - \tau_{lt})w_t a_t \bar{l}_d + T_t \\
(3) \quad & k_{pt+1} = [(1 - \delta)k_{pt} + i_{pt}]/(1 + \gamma_n) \\
(4) \quad & N_t = (1 + \gamma_n)^t \\
(5) \quad & c_{ct}, c_{dt}, i_{pt} \geq 0,
\end{aligned}$$

where k_{pt} is the beginning-of-period capital stock in period t , r_{pt} is the rental rate paid for that capital, w_t is the wage rate in t , τ_{kt} and τ_{lt} are proportional tax rates on capital income and labor income, respectively, in t , $R_t b_t$ is the value of matured government debt, and T_t are government transfers.⁸ All quantities are in per capita terms; the constant growth rate of the population N_t is given by γ_n . The processes for a_t , r_{pt} , w_t , τ_{kt} , τ_{lt} , R_t , and T_t are viewed exogenously by the family and are specified later.

Distinguishing between workers and hours per worker is required for our robustness analysis that extends the benchmark model to include variable capacity utilization. Moreover, this specification is useful because the marginal disutility of working for those in the army does not affect the choices made by civilians. This is important because there is considerable uncertainty over this marginal disutility of labor for the draftees, as there is limited information on the number of hours worked by those in the military during this period, and the disutility of military work probably differs from private work. Thus, any errors in measuring the work disutility of draftees will not affect the private choices in our formulation, which is very similar to that in Hansen (1985), Kydland and Prescott (1991), Cole and Ohanian (2002), and Hayashi and Prescott (2002), all of which distinguish between workers and hours per worker.⁹

There is a single physical good that is produced from a constant returns to scale technology. The technology is operated by a competitive representative firm, which hires private capital, public capital, and labor services. Output, which we measure as nonmilitary output, is given by

$$(6) \quad Y_t = F(K_{pt}, K_{gt}, Z_t L_{pt}),$$

where K_{pt} is the beginning-of-period private capital stock for the economy in t , K_{gt} is the beginning-of-period public capital stock used by the private sector in t , Z_t is the level of labor-augmenting technology in t , and L_{pt} is the total labor input in the nonmilitary sector in t . We assume that the

level of labor-augmenting technology is equal to $z_t(1 + \gamma_z)^t$, where z_t is a productivity shock.

We include government capital in production because the federal government financed increases in industrial construction and producers' durable equipment during World War II, including significant investments in the aircraft, automotive, and aluminum industries. Gordon (1969) estimates that government-owned, privately operated capital increased the manufacturing capital stock by 30 percent between 1940 and 1945. (See also Gordon (1970), Jaszi (1970), and Wasson, Musgrave, and Harkins (1970).) We denote government investment expenditures by I_g .

Government purchases of consumption goods are denoted by C_g , and government payments to military personnel are denoted by $N_t w_t a_t \bar{l}_d$. Total government spending is the sum of the three expenditure items:

$$(7) \quad G_t = C_{gt} + I_{gt} + N_t w_t a_t \bar{l}_d.$$

Government capital evolves according to the following law of motion:

$$(8) \quad K_{gt+1} = (1 - \delta)K_{gt} + I_{gt}$$

with K_{g0} and the process for I_{gt} given. We assume that private and public capital depreciate at the same rate δ . We also assume that the government satisfies the present value budget balance. The government budget constraint is given by

$$(9) \quad B_{t+1} = G_t + R_t B_t - \tau_{lt} N_t w_t ((1 - a_t) l_{ct} + a_t \bar{l}_d) - \tau_{kt} (r_{pt} - \delta) K_{pt} - r_{gt} K_{gt} + T_t.$$

We close the model by specifying the functions that the family treats exogenously when solving its optimization problem in (1). Since firms are competitive, the rental prices for the factors of production are equal to their marginal products. Therefore, the rental rates in (2) and (9) and the wage rate in (2) are equal to the partial derivatives of the production function F in (6) with respect to K_p , K_g , and L_p , respectively. Government debt that is accumulated during the war is retired following the war. The technical appendix (McGrattan and Ohanian 2008) shows that the results are not sensitive to alternative specifications of postwar tax rates that retire this wartime debt.

There are six exogenous variables in the model: conscription (a_t), the tax rate on capital income (τ_{kt}), the tax rate on labor income (τ_{lt}), government consumption (C_{gt}), government investment (I_{gt}), and labor-augmenting productivity (z_t). The evolution of the six exogenous variables

is governed by a state variable, s_t , which specifies a particular set of values for a_t , τ_{lt} , τ_{kt} , C_{gt} , I_{gt} , and z_t .

An *equilibrium* for this economy consists of the following: allocations for households c_{ct} , c_{dt} , l_{ct} , i_{pt} , and k_{pt} ; inputs for firms K_{pt} , K_{gt} , and L_{pt} ; and sequences of prices r_{pt} , r_{gt} , w_t , and R_t that satisfy the following conditions: (i) taking prices and exogenous policies for a_t , τ_{kt} , and τ_{lt} as given, households maximize utility subject to constraints (2)–(5); (ii) taking prices as given, firms maximize profits period by period $Y - r_p K_p - r_g K_g - w L_p$; (iii) factor markets clear:

$$(10) \quad K_{pt} = N_t k_{pt}$$

$$(11) \quad L_{pt} = N_t(1 - a_t)l_{ct};$$

(iv) the resource constraint

$$(12) \quad C_{pt} + I_{pt} + C_{gt} + I_{gt} = Y_t$$

holds, where $C_{pt} = N_t[a_t c_{ct} + (1 - a_t)c_{dt}]$ and $I_{pt} = N_t i_{pt}$; and (v) (9) is satisfied.

To test the robustness of our results, we will also consider a version of the model in which there is variable capacity utilization. We will see that the results are not sensitive to this modification, and therefore we present the variable capacity model in McGrattan and Ohanian (2008).

Table 1 summarizes the values of all parameters, which we discuss in detail below. Preferences are given by

$$(13) \quad U(c, l) = \log(c) + \psi(1 - l)^\xi / \xi,$$

which implies a compensated labor supply elasticity of $(1 - l)/[l(1 - \xi)]$. We choose a benchmark value of $\xi = 0$, which implies log preferences over leisure. In McGrattan and Ohanian (2008), we evaluate the robustness of our results by choosing a value of ξ that yields a lower labor supply elasticity.

The parameter ψ is chosen so that the fraction of time allocated to nonmilitary work in the deterministic steady state is 26.6 percent, which is consistent with the observed U.S. average over the period 1946–1960. In principle, we also need to specify the exogenous hours requirement for those in the military. We use 50 hours per week, which lies between estimates of average hours per week for civilian workers in the federal government (see Kendrick 1961) and average hours per

week for military workers in basic training (see Siu 2008). This choice does not affect our results, since the level of military hours does not impact private choices.

We assume that government capital and private capital are perfect substitutes. As discussed above, this is a reasonable assumption, as most of this government investment was in government-owned, privately operated plant and equipment. The technology is given by

$$(14) \quad F(k_p, k_g, zl) = (k_p + k_g)^\theta (zl)^{1-\theta},$$

implying that TFP is equal to $z^{1-\theta}$. We chose $\theta = 0.38$, which is consistent with the U.S. share of income paid to capital during this period. The parameter β is chosen so that the capital-output ratio is consistent with the U.S. level during the war. The depreciation rate (δ) for both government and private capital is 5.5 percent.

The growth rates of trend labor-augmenting technological progress and the population are set to their average values over this period: $\gamma_z = 2.0$ percent and $\gamma_n = 1.5$ percent.

For our baseline results, we assume that households have perfect foresight about the paths of exogenous variables in Figure 1. For years beyond 1946, we assume that the values in 1946 persist. In McGrattan and Ohanian (2008) we analyze stochastic simulations of our model and alternative scenarios for postwar expectations. We also show how to compute the equilibrium decision functions for our model in both the deterministic and stochastic cases using a finite element method. (See also McGrattan 1996.)

6. Comparing the Model to the Data

We compare seven variables from the model to their counterparts in the U.S. data. Real GNP, consumption, investment, two measures of hours worked, and two measures of factor productivity are compared to U.S. counterparts derived from the national income and product accounts (U.S. Department of Commerce 1986), the reproducible tangible wealth tables (U.S. Department of Commerce 1987), and data on manhours (Kendrick 1961).

We compare production plus military compensation in the model ($Y + wNa\bar{l}_d$) to real U.S. GNP. We compare private consumption in the model to U.S. personal consumption expenditures on nondurables and services, and we compare private investment in the model to U.S. gross private domestic investment plus durable consumption expenditures. We subtract indirect business taxes for sales from GNP and consumption expenditures. We impute a service flow for durables

equal, in real terms, to 4 percent times the stock of durables.¹⁰ For both the model and the data, we report GNP, consumption, and investment in per capita detrended terms as we did for the government spending series in Figure 1. Specifically, we divide the series by the population over 16 and by the growth rate of labor-augmenting technology.

We compare per capita nonmilitary hours worked in the model to nonmilitary hours from Kendrick (1961) divided by the population over 16. For completeness, we compare total per capita hours—nonmilitary plus military—in the model $((1-a)l_c + a\bar{l}_d)$ to total U.S. manhours in Kendrick (1961) divided by the population over 16. Both the actual and model per capita hours series are normalized by discretionary time, which is 12 hours per day.

Finally, we compare measures of factor productivities in the model and data. For capital, we compare nonmilitary output Y divided by total capital $K_p + K_g$ to its counterpart in U.S. data. The output and capital measures are the same as that used in computing TFP. To put this ratio in more interpretable units—that is, in the units of an after-tax return—we multiply both the U.S. and model series by the capital share times one minus the tax rate on capital $(\theta(1 - \tau_k))$ and then subtract the depreciation rate δ . For labor, we compare nonmilitary output Y divided by nonmilitary hours L_p multiplied by $(1 - \theta)(1 - \tau_l)$ to its counterpart in U.S. data. In the appendix, we also discuss several alternative measures of factor returns and how they compare to the model’s predictions.

We calculate the equilibrium of the model beginning in 1939, given initial values for private and public capital, which we take from the U.S. Department of Commerce (1987). We compare the results to the data by plotting the actual realizations of output, consumption, investment, labor, and factor productivities between 1939 and 1946 against the model realizations for these variables.

A. Results

Figures 2–4 show the results.¹¹ The lines with open squares are the U.S. series, and the lines with the filled circles are the model series. The main finding is that the time series for the model and data are quite similar. The model captures the magnitude of the changes that occur in these variables, and also captures much of the timing of these changes as well.

The first plot in Figure 2 is real detrended GNP. Both model and data GDP rise about 40 percent over the course of the war, and both decline significantly after 1945. The second plot shows private consumption in both the model and the data. Model consumption shows almost no change, while actual consumption shows a decline of about 2 percent between 1939 and 1944. The third

plot in Figure 2 shows that private investment in both the model and the data declines significantly through 1944, and then recovers after government spending declines after 1945. Thus, the model reproduces several main features of actual wartime changes in output and its components.

Figure 3 plots nonmilitary hours, which is the focus of the analysis, as well as total hours. All of these series are divided by the 1946–1960 U.S. average fraction of time at work to provide hours measures that are relative to a postwar trend. Both the data and the model show a significant increase in hours during the war, followed by a decline. The main deviation between the model and the data is in terms of timing; the model hours initially rise more quickly than actual hours, and model hours begin declining one year before actual hours decline. Part of the reason for this deviation is our choice of perfect foresight. In McGrattan and Ohanian (2008), we relax this assumption but still find that hours rise too much initially relative to U.S. hours because households expect higher future tax rates on labor.

Figure 4 shows measures of after-tax factor returns. The first plot shows the time series for the after-tax marginal product of capital (as a percentage). The model predicts a return that is roughly 3.4 percent for the first half of the war, rising to about 3.9 percent in 1944, then declining afterward. The actual return is quite similar, averaging 3.4 percent early in the war, rising to about 4 percent, and then declining afterward.

The second plot in Figure 4 is the after-tax return to labor. As before, nonmilitary hours are normalized; they are divided by the 1946–1960 average fraction of discretionary time in work. During the war, the model shows a slight increase in the after-tax return to labor of about 2.8 percent between 1941 and 1944, whereas the actual return shows a slight decrease of -0.6 percent. But, as we show later, the prediction of little change in the after-tax return to labor is consistent with most U.S. measures of this return. After the war, both returns show a steep decline.

B. Discussion

To understand the economic forces driving these results, we present the relative contributions of each exogenous variable. We do this by computing the equilibrium of the model for the following set of experiments: the first experiment has only government spending, as the values of all other exogenous variables are held fixed at their 1939 values. The second experiment adds to (1) productivity shocks, the third adds to (2) the draft, and the fourth adds to (3) distorting taxes, which is the full model.¹² We compare results across these experiments to see how the addition of each variable impacts the results.

Tables 2 and 3 show the results of these four experiments for the two labor market variables, hours worked and the after-tax wage. We choose these two variables for this detailed analysis because they have attracted the most attention in the literature. The experiments show that government spending has the largest impact on the economy, that all the shocks have substantial effects, and that consequently all shocks are required for the model to match the data as closely as it does.

Table 2 shows that government purchases have the largest quantitative impact, as they increase nonmilitary hours by nearly 30 percent on average during the period 1939–1946. That is, if the war only involved higher government spending, then the theory predicts that hours should have increased about four times as much as observed. Instead, the reason that model hours are on average similar to actual hours is largely due to higher labor income taxes, which reduce hours by as much as 16 percentage points at the peak of the war, and to the draft, which reduces hours by as much as 6 percentage points.

Thus, the key reason in the model that nonmilitary hours rise is because of the enormous resource cost of the war. This effect is attenuated somewhat by higher taxes, which reduces the incentive to work, and by the draft, which reduces the household's time endowment. Productivity tends to reduce hours early, as households foresee higher future productivity, and it then generates somewhat higher hours toward the end of the war as productivity peaks.

Table 3 shows the results for after-tax wages. Again, government spending has the largest impact in the model. In the absence of other shocks, the wartime expenditure shock drives down the after-tax wage 20 percent below trend by the end of the war. As in Table 2, higher taxes and the draft offset some of the impact of higher government spending. Productivity has a large impact on wages, increasing wages by about 15 to 20 percent at the end of the war.

Table 3 also compares the model after-tax wage to a number of after-tax wage measures from the data. It is useful to make a broad comparison across alternative wage measures since the data may have some measurement error, and because different wage series have been used in the literature.¹³

Most of the empirical measures are similar to the model wage. Labor productivity rises by 1 percent, and its average is 1 percent above trend. The wage of all civilian workers rises by about 8 percent, and its average is on trend. The nonfarm wage initially declines, and then rises. Its mean is 4 percent above trend. The manufacturing wage declines by about 8 percent, but its average is

5 percent above trend. Recall that the model wage rises about 6 percent between 1939 and 1943, to roughly 3 percent above trend, and its average value during the war is on trend.

The after-tax manufacturing wage differs from the other wage measures. It falls the most over the war, and despite this decline, its average is also the highest above trend. These features of the manufacturing wage are likely due to New Deal labor policies, which significantly impacted manufacturing. Specifically, manufacturing wages rose considerably above trend in the 1930s, as a consequence of the National Industrial Recovery Act (1933) and the National Labor Relations Act (1935), both of which increased labor bargaining power (see Cole and Ohanian (2004)). However, real manufacturing wages then began declining during the war, reflecting policy changes that reduced union bargaining power, including the National War Labor Board, which reduced negotiated wage increases in several industries (see Cole and Ohanian (2004) and Atelson (1998)).

Given that manufacturing wages were impacted by changes in unionization policies before and during the war, it is difficult to use the manufacturing wage, or changes in that wage, to measure the changes in the incentive to work. New Deal labor policies had a smaller impact on broader measures of wages, such as the other three measures in the table. Consequently, the broader wage measures, with changes that range from a flat profile over the war (labor productivity) to an increasing wage profile (wage for all civilians), are better comparisons to the model wage.

In summary, this comparison shows that both model hours and wages are similar to their data analogues. The following section presents additional analysis of hours and wages to address Mulligan's (1998) view that higher hours are a puzzle during the war. It also addresses other reservations about the model that have been expressed in the literature.

7. Addressing the Critiques

Section 2 presented a number of critiques of whether the model can successfully account for the impact of World War II. We now address those critiques.

A. Analyzing After-Tax Wages and Hours

The main critiques concern the neoclassical model's ability to account for patterns in after-tax wages and hours. We begin with Mulligan's critique (1998). His analysis includes two of the wage series we use (civilian and manufacturing), and concludes that the after-tax real wages were not sufficiently high to induce higher hours: "Empirical support . . . cannot be found because after-tax wages do not appear to be temporarily high during the war period. The primary force working

against wage motives is the massive across-the-board income tax increases that occurred during the war” (p. 1071).¹⁴

To address this issue, we assess how much wages would have to rise in the model to generate the observed increase in nonmilitary hours, and we then compare that model wage to observed wages. To do this, we consider that there may be measurement error in the available tax rate estimates, and we construct a sequence of tax rates such that hours in the model between 1939 and 1946 are equal to actual hours:

$$(15) \quad l_{pt}^m(\hat{\tau}_t) = l_{pt}^d, \quad t = 1939, \dots, 1946,$$

where l_{pt}^m is per capita nonmilitary hours in the model, $\hat{\tau}_t$ is the constructed tax rate, and l_{pt}^d is actual nonmilitary hours divided by the population over 16.

Figure 5 shows the model wage that generates the observed increase in military hours (which we call the “model implied wage”) and also shows the four empirical wage measures: the manufacturing wage, the nonfarm wage, labor productivity, and the civilian wage. These are all pre-tax measures since we are starting from the premise that tax rates may be mismeasured. The empirical measures are all detrended by the growth in technology and reported relative to their respective averages between 1946 and 1960.

To match the path of U.S. nonmilitary hours, we need a tax rate series that averages about 15 percent and rises less over the period 1939–1946 than the Joines’ (1981) tax rate series. (See McGrattan and Ohanian (2008) for a graph of alternative measures of the labor tax rate.) If taxes are rising between 1939 and 1946 and households perfectly anticipate higher labor tax rates in the future, then they increase their hours early in the period to take advantage of the low tax rates. During the first few years, TFP is also relatively low, but it is a less important factor for labor supply than the tax on labor.

In terms of comparing percentage increases in the wage series over the war, the implied model wage series *rises less* than the most comparable empirical wage measures. The implied model wage rises about 12 percent between 1939 and 1945 compared to increases of about 16 percent for actual labor productivity, about 14 percent for the nonfarm wage, and about 22 percent for the civilian wage. It is virtually identical to actual labor productivity between 1939 and 1942, and remains within 2 percentage points between 1943 and 1946. The implied model wage lies between the other wages series, as it is below the nonfarm wage and manufacturing wage, and is above the civilian wage through 1942. By 1943, all of the empirical measures are above the implied model wage.

The implied model wage increase is higher than the manufacturing wage, which increases about 6 percent, though as we noted earlier, manufacturing is probably the least informative of these wage measures due to the fact that manufacturing wages were impacted by changes in New Deal labor policies. Note that the average percentage increase across all four wage measures is 12 percent, which is the percentage increase in the model wage for this experiment. This suggests that the observed increase in nonmilitary hours is consistent with observed changes in wages.

These results raise the question of how much would labor tax rates have to rise in order to eliminate higher hours worked. To address this question, we calculated the model equilibrium with labor tax rates such that hours worked on average were equal to their trend value. We find that tax rates would have to be much larger than those reported by Joines (1981)—or any other estimates of U.S. tax rates—to prevent nonmilitary hours from rising in the model. The tax rates would have had to jump to 25 percent at the start of the war and continue to rise to close to 30 percent by 1943. (See McGrattan and Ohanian 2008.) When we rerun the experiment varying capital tax rates, we find that even when they are set to 100 percent during the war, hours rise.

The main reason why hours rise so much is because of the enormous government spending increase. The present value of war expenditures, measured as the difference between actual government expenditures between 1939 and 1945 and peacetime government expenditures, is 180 percent of steady state GDP. This large reduction in wealth induces higher hours. Note that at the peak of the war in 1944, total government spending was a little over 80 percent of trend output. This means that if labor *had not increased*, private consumption would have dropped substantially. Households in the model respond to this large resource drain by working more.

B. Other Critiques

Regarding other critiques, Rotemberg and Woodford (1992) raise questions about the neoclassical model's ability to account for pre-tax real wages during wars and during other periods of large exogenous increases in government spending. The average percent deviation between the pre-tax marginal product of labor in the perfect foresight model and in the data between 1941 and 1946 is $-.2$ percent, with the largest annual percent deviation during this period equal to 2 percent. As we see from Figure 5, with the exception of the manufacturing wage rate, the other measures are close to the marginal product of labor.¹⁵

Blanchard and Perotti (2002) have questioned the ability of the theory to account for the division of output between consumption and investment in response to large fiscal shocks. In 1944,

private consumption's share of GNP in the perfect foresight model is 40 percent, compared to 37 percent in U.S. data. In 1944, private investment's share of GNP in the perfect foresight model is 2 percent, compared to 6 percent in U.S. data.

Burnside, Eichenbaum, and Fisher (2004) find that the neoclassical model requires additional features, including habit formation in preferences and investment adjustment costs, to account for movements in labor during periods of post-World War II military buildups. Resolving these differences is beyond the scope of this study, as the methodologies and the episodes studied differ considerably in the two papers. We therefore leave this for future research.

8. Sensitivity Analysis

In McGrattan and Ohanian (2008) we present detailed results on additional experiments that further assess the robustness of our results. First, we evaluate how low the labor supply elasticity would have to be such that, on average, nonmilitary hours did not rise during the war. Second, we introduce variable capacity utilization in our model, allowing for variation in the number of employed and the number of hours in the workweek, since both rose significantly during the war. Third, we conduct different experiments varying household expectations about the exogenous variables, altering the timing of when households learn about the war, what they expect to happen during the war, and what they expect after the war. Here, we briefly describe the results for each of these experiments; further details can be found in McGrattan and Ohanian (2008).

To keep the average level of nonmilitary hours equal to its trend, we need the labor supply elasticity to drop by more than a factor of eight relative to the log utility case used in the benchmark model. The elasticity for the log utility case is 2.75 percent compared to 0.32 percent in this experiment. This latter elasticity is much too low for an aggregate representative household model. To see this, consider replacing the log utility function in the prototype business cycle models studied by McGrattan (1994) with (13). We can set ψ and ξ so as to achieve the same steady-state hours worked and lower labor elasticities. In her benchmark case with technology shocks only and divisible labor, a labor elasticity of 0.5 generates a standard deviation of hours worked equal to 0.3; the standard deviation for U.S. hours is 1.52. (See McGrattan 1994, Table 1.) Similar results are found for her model with taxes. In that case, a labor elasticity of 0.5 generates a standard deviation of hours worked equal to 0.51—again, much lower than that in the data. For a labor elasticity of 0.32, the results are even more striking: the standard deviations of hours worked predicted by the model are in the range of 0.22 to 0.38—significantly below the data. This implies that implausibly

low aggregate labor supply elasticities are required in our model to choke off the World War II economic expansion.

Our second experiment involves adding variable capacity, along the lines of Kydland and Prescott (1991). In this case, the technology is given by:

$$(16) \quad Y_t = (h_t K_t)^\theta (h_t Z_t N_t)^{1-\theta},$$

in which h is the length of the workweek, N is private employment, K is the sum of public and private capital, and Y is nonmilitary output. To induce endogenous variation in the workweek, we assume individual preferences are given by

$$(17) \quad U(c, h) = \log(c) + V(1 - h) - \eta\chi\{h > 0\},$$

where η measures the disutility of entering the workforce and χ is an indicator function. Aggregating utility costs of entering the workforce across individuals yields an aggregate utility function that depends on both the workweek and the number employed.

In terms of inputs in Figure 1, all remain the same with the variable capacity utilization model except TFP, which is derived from (16). The time path of TFP does differ from TFP in Figure 1 for the benchmark case, but the response of the economy does not. Between 1941 and 1944, the TFP increase in the variable workweek model is about 10 percent, compared to an increase of about 13 percent in the benchmark model. The responses of the endogenous variables in the variable capacity model, however, are very similar to those from the benchmark model, as the smaller increase in TFP in the variable workweek model is offset by a larger impact of a TFP change on the endogenous variables. For example, GNP in 1943 is 36 percent above trend in the benchmark model and 33 percent above trend in the variable capacity utilization model. Nonmilitary hours in 1943 are 11 percent above trend in the benchmark model and 9.4 percent in the variable capacity utilization model. Because of the similarity, we report the results side-by-side in McGrattan and Ohanian (2008).

When altering household expectations, we tried several different experiments. We started the analysis in 1941 rather than 1939. The change in timing of when the households learn about the war hardly changes the results. We allowed for uncertainty about entering the war and its duration, choosing transition matrices for the exogenous variables that generate time series consistent with earlier U.S. wars. The main differences between the stochastic and deterministic results is the pattern of consumption and nonmilitary hours. With uncertainty, the model generates some decline

in consumption during the war and lower nonmilitary hours prior to the start of the war. But, the overall quantitative fit of the model is very similar for the deterministic and stochastic versions of the model. When we change expectations about the exogenous variables in the postwar period, we do generate large differences relative to the benchmark if we assume households expect TFP or tax rates to fall back to levels seen in the Great Depression. However, if all fiscal states, including government spending, are expected to fall back to 1930s levels while TFP is expected to remain high, then we find results similar to the benchmark case.

In summary, our sensitivity analysis shows that the results are fairly robust to changes in the labor supply elasticity, to the inclusion of variable capacity utilization, and to varying expectations about the timing and duration of the war and to plausible expectations about the postwar.

9. Summary and Conclusion

The behavior of the World War II economy was well outside normal bounds. Military expenditures increased 500 percent, tax rates increased as much as 100 percent, the draft increased 500 percent, there was rationing, price and wage controls, and interest rate control of the public debt, private investment fell nearly to zero, and output roughly doubled. There is no consensus theoretical framework for understanding this important period, largely because there has been no detailed quantitative assessment of alternative models in response to a fiscal shock, and policy response, of this size.

This paper conducted the most comprehensive evaluation of the growth model to date in response to World War II. We found that the behavior of the World War II economy was similar to the neoclassical growth model simulated in response to wartime shocks. Specifically, the time paths of output, consumption, investment, labor input, wage rates, and returns to capital from the model in response to World War II government spending, the draft, tax rates, and productivity shocks are quite similar to the data.

We conclude that the growth model, at least in the case of World War II, can reasonably capture the quantitative impact of very large fiscal shocks. The results suggest several avenues for future research. One is based on the interesting implication that the many wartime regulations and restrictions adopted during World War II do not appear to significantly impede the ability of the model to account for several aggregate variables during this period. This suggests that an important topic is to understand *why* these restrictions did not seem to have important effects on the aggregate economy. Another topic to investigate is the relatively low observed labor input

at the early stages of the war, and the extent to which that is attributable to New Deal labor policies that depressed employment, as suggested by Cole and Ohanian (2004). A third topic is understanding the contribution of R&D, the development of operations research and management sciences practices, and other factors to the wartime productivity increase. Finally, an international comparison would provide useful variation between countries that sustained battle activity within their borders (United Kingdom, France, Japan, Germany) and countries in which most of the war was outside their borders (United States).

Appendix. Alternative Measures of Factor Returns

In Section 6, we used marginal products constructed from U.S. data to assess the model's predictions for returns to capital and labor. In theory, we can also use factor incomes per unit of input and, in the case of capital, returns on assets held by capital owners. In this Appendix, we discuss these alternatives and relate them to the measures we used above. For labor, we find that the marginal product from the model accords well with measures of compensation per hour. For capital, we find that the marginal product on average is consistent with returns based on capital incomes and stocks and with equity returns. We also explain why another measure of returns—based on debt assets—is not the appropriate measure for our model economy because of regulations and restrictions impacting the bond markets.

A. Return to Labor

In Figure 4, we compared measures of labor productivity (after tax) for the model and U.S. data. Figure 5 plots the pre-tax labor productivity in the United States based on U.S. national product, along with three measures of labor based on U.S. national income.^{16,17} All compensation rates are divided first by the deflator for GNP less military compensation and then by the growth trend for technology.

The two broadest measures of aggregate real compensation per hour—nonfarm and civilian—are very similar to the marginal product of labor from the model. These compensation measures are the most reasonable comparison to the model marginal product because they are economy-wide wage measures and because they are much less affected by changes in union bargaining power than is manufacturing compensation per hour. In particular, there were substantial declines in union bargaining power during World War II. Cole and Ohanian (2004) argue that manufacturing wages rose sharply in the mid- to late 1930s as a consequence of large increases in union bargaining power, and that this bargaining power declined sharply during the 1940s. During the war, wages were no longer set by collective bargaining, but rather by the National War Labor Board (NWLB), which routinely rejected negotiated wage settlements between firms and unions. After the war, the Taft-Hartley Act further reduced union bargaining power. Cole and Ohanian estimate that most of the increase in manufacturing wages generated by unions/cartels in the 1930s had vanished by 1947 as a result of the NWLB decisions. This finding is consistent with the fact that pre-tax manufacturing

compensation per hour did not increase much during the war. (See Figure 5.) It is reasonable to expect that introducing union bargaining, and distinguishing between manufacturing and non-manufacturing sectors, would allow the model to account for differences between manufacturing and non-manufacturing compensation. However, this is well beyond the scope of this analysis.

B. Return to Capital

In Figure 4, we used the after-tax marginal product of capital when comparing the return to capital in the data and model. Here, we discuss three alternative measures for the empirical counterpart of our model's return: a measure of the return on capital found by dividing after-tax capital income by the capital stock, a measure of the return on equities, and a measure of the return on bonds.

First, we can compare the model's return on capital to some measure of capital income divided by the appropriate stock—both nominal since the Bureau of Economic Analysis (BEA) does not report real incomes. Before making these comparisons, we need to address several issues that arise when doing these calculations. One issue is data revisions: there have been significant revisions in nominal stock estimates across BEA reports, especially for structures. For example, for the period 1941–1946, the BEA estimates show an average level for the current-cost net stock of private nonresidential structures of \$126.4 billion in the U.S. Department of Commerce (2002) and \$70.4 billion in U.S. Department of Commerce (1987). Another issue that arises, especially in the case of World War II, is the possibility that factor payments by government to business are not allocative period-by-period. For example, in cases where the capital was government-owned but privately operated, many contracts prespecified the returns on the investment. (See Braun and McGrattan 1993.) If capital returns are not allocative, then we can only compare average marginal products with average per-unit incomes. A related issue is how to attribute returns to assets, such as government assets, whose incomes are not included or imputed in the national income and product accounts (NIPA).

McGrattan and Prescott (2003), using recent BEA data, estimate an after-tax return to noncorporate capital of 4 percent over the period 1929–2000; over the period 1941–1946, the average return for their series is 4.95 percent. They impute a 4 percent return to government capital, which may be high for the World War II period. They also leave out the corporate sector because of issues with estimating returns to intangible capital. If we redo their calculation imputing a 0 percent return to government capital and adding in measured corporate income less tax (in the

numerator) and measured corporate capital (in the denominator), then the average we compute is 4.5 percent over the period 1941–1946. As McGrattan and Prescott point out, however, this is an overestimate of the true return because measured corporate income does include part of the income to intangible capital, while measured corporate capital does not. This puts the estimate of the average closer to the average of the after-tax marginal product of capital, which is 3.55 percent for the U.S. data and 3.52 percent for our benchmark model.

Another possible comparison that can be made is the model’s return on capital and the return on corporate equities. When comparing the marginal product of capital with any equity return, the main deviation is in the volatility rather than the means. For example, averaging the inflation-adjusted total returns for the Standard and Poor’s (S&P) Composite reported in the Ibbotson Associates (2001) yearbook for the period 1941–1946, we find an average of 6.87 percent. If we adjust for taxes on dividends using McGrattan and Prescott’s (2003) tax rate on dividends times the S&P income return from Ibbotson Associates, then the average return is 4.1 percent. This slightly overestimates the actual equity return because of taxes on capital gains; however, these adjustments are difficult to calculate given the fact that only realized gains are taxed. If we abstract from taxes on capital gains, the average equity return is only 50 basis points higher than the model’s predicted return.

McGrattan and Prescott (2003, Figures 3–5) find that differences in average returns on bonds, stocks, and NIPA capital are not large *except* in the case of debt returns during the wartime period that includes World War II and the Korean War. Specifically, averaging inflation-adjusted total returns reported in the Ibbotson Associates (2001) yearbook for the period 1941–1946, we find corporate bonds earned –3.7 percent, long-term government bonds earned –3.5 percent, intermediate-term government bonds earned –5.1 percent, and U.S. Treasury bills earned –6.4 percent. These very low numbers for debt returns have been noted in earlier work as puzzling for neoclassical theory. (See, for example, Mulligan 1998.)

But these returns are not a good measure for the household intertemporal marginal rate of substitution, which is the relevant object in the model. Federal government debt made up only about 5 percent of total household assets in 1945. Most federal debt was held by the Federal Reserve System, commercial banks, insurance companies, pension funds, state and local governments, and government retirement accounts. Almost all of these institutions were restricted in terms of the assets they could hold; they typically could not hold equity, and their portfolios were held primarily in credit market instruments (debt) or cash. For example, life insurance companies, banks, savings

institutions, pension funds, and state and local government retirement programs held between 50 and 65 percent of their assets in U.S. Treasury securities in 1945, compared to only about 5 percent of household assets in federal government debt. The literature on financial markets discusses the regulations and restrictions in detail, and their impact on portfolios. We have reproduced some passages from that literature below.

“There have been several factors influencing the amount of stocks held by the companies: (1) statutory requirements and limitations, (2) valuation regulations, and (3) investment policies. With respect to statutory provisions . . . the provisions of New York State law in particular have had an important influence on the amount of stocks held by all life insurance companies since companies domiciled in the state have constituted a large, although varying, segment of the industry. Under the New York law, investment in preferred stocks was not permitted from 1906 until 1928 and in common stocks from 1906 until 1951. The laws of other states also have not permitted stock investment. When such investment has been permitted there have been limitations on the amount of such holdings. These limitations are often expressed as a percentage of assets or surplus and help to account for the relatively small investment in equities by life insurance companies.” Life Insurance Association of America (1962, p. 51)

“Although commercial banks are the largest single group of financial institutions if measured by size of assets, they have hardly ever been important holders of corporate stock. This fact is mostly due to regulation. National banks are virtually precluded from owning corporate stock except that of Federal Reserve banks. While the regulations are not as strict in many states, they still severely limit the freedom of state-chartered banks to invest in corporate stocks even if they desire to.” Goldsmith (1973, p. 52 and Table 2.7)

“In the years immediately following the war [WWII], the interest rates on long-term government bonds (pegged at 2 1/2 percent) kept interest rates on private bonds at similar low levels. The higher returns on common stock investments were strong inducement for bank trustees to invest an increasing share in stocks. Accordingly, uninsured pension funds quickly sold off the government securities which they had accumulated during WWII and invested primarily in corporate stocks and bonds. This change was made possible by a revision in New York State law allowing trustees to invest up to 35 percent of a fund in stocks.” (p. 230) “Historically, life insurance companies have been very conservative investors, on the presumption that their fundamental objective should be safety of principal. As a result over three-fourths of all life insurance assets have been invested in corporate bonds and mortgages. A variety of statutory and institutional considerations reduced the investment alternatives in corporate stock that were available to life insurance companies; state laws provide very strict limitations. Most life insurance company assets are held by companies licensed in New York. Originally, New York State law prohibited investment in corporate stock. Relaxation of this restriction in 1951 allowed life insurance companies to invest up to 3 percent of total assets in common stock; an amendment in 1957 raised the limit to 5 percent. The law also prescribes limits on the type of company whose stock is eligible. A company must have paid a dividend in each of the previous ten years, and dividends must not have exceeded earnings in any year. Obviously, these restrictions severely limit the choice of stocks open to life insurance companies.” (p. 231) “The rules for valuation of assets constitute the second major deterrent to stock investment by life insurance companies. Most life insurance companies are mutual companies and are required by law to return profits in excess of

a stated level of net policy liabilities. Thus, determining asset values critically affects a company's cash flow and almost since its beginning has been the subject of dispute in the industry." (p. 253) "A trust agreement is an arrangement by which the trustee assumes fiduciary responsibility for managing assets for the benefit of another. The agreement typically defines that responsibility, the degree of discretion of the trustee, and the rules for distributing benefits of the trust. The definition of fiduciary discretion has many dimensions. Often it limits the extent of corporate stock and other types of investments; it may impose limits on the share of funds that may be invested in a single company; and it may lay out guidelines, indicating which companies are eligible. Also state laws and state courts interpret the nature and limits of trustee discretion differently. In some cases the trustee is limited to selecting from a 'legal list' of eligible investments maintained by many states. Within the agreed upon limits of fiduciary responsibility trustees typically are limited by the 'prudent man' rule. (Harvard College vs. Amory 1835)." (Goldsmith, Brady, and Mendershausen 1956, p. 226)

These passages summarize why federal debt was almost exclusively held by these institutions. Of the very small amount of federal debt that was held by households, roughly two-thirds was held in savings bonds. These savings bonds had their own set of idiosyncratic attributes: there was no risk of capital loss, owners were insured against loss or theft, the securities were non-marketable, acquisition costs were very low (many savings bonds were acquired through payroll deduction), they were available in small denominations (as low as \$25), and interest rates on some of these securities were higher than on long-term Treasury securities. Savings bonds were primarily accumulated by relatively low-income households who found the low denominations and low acquisition costs attractive, and who in fact typically paid no taxes on the income from these securities. Low-income households at this time had few other investment options because the cost of holding a diversified equity portfolio was high, and minimum denominations for other forms of debt were typically much higher than for savings bonds. (See Board of Governors 1944.) This discussion indicates that analyzing U.S. debt returns during World War II would require modeling the details of regulations affecting the bond markets, which is well beyond the scope of this paper.

In summary, we have shown that the model's predicted wage rate is consistent with standard measures of U.S. labor productivity and compensation per hour. We have shown that the model's predicted return to capital is consistent with a measure of the return to NIPA capital, based on the marginal product of capital; it is consistent on average with a measure of the return based on the ratio of capital income to the capital stock and on average with a measure of the return to corporate equities.

Notes

¹By the *neoclassical model*, we mean the one-sector optimal growth model and a constant returns to scale technology with capital and labor as inputs, a standard law of motion for the capital stock, balanced growth preferences defined over consumption and leisure, a resource constraint that divides output between consumption, investment, and government spending, and perfect competition in all markets.

²A few studies have used the neoclassical model to address a limited set of questions about World War II. Braun and McGrattan (1993) use a stochastic model with World War II government expenditure shocks and focus on whether the model is consistent with pre-tax real wage changes during the war. However, their analysis omits all other shocks and does not provide a systematic test of the other variables in the model. Ohanian (1997) uses a perfect foresight neoclassical model for normative rather than positive purposes. He measures the welfare costs of the different war finance policies used in World War II and the Korean War. His analysis does not shed light on the present disagreement about the appropriate theoretical framework for analyzing large fiscal shocks.

³Rotemberg and Woodford (1992) analyze data on military expenditures and real wages from World War II, and also from other years. The World War II observations are likely the most informative because they are by far the largest expenditure shocks and also the most exogenous of the increases in military spending. They do not perform an assessment of the model for World War II.

⁴Siu (2008) studies the welfare consequences of World War II conscription in a Ramsey environment in which taxing the household's time endowment is one component of efficient taxation.

⁵Most of these R&D expenditures were outside of the Department of Defense. Real defense spending rose from \$30 million in 1930 to \$425 million in 1945, which implies that about two-thirds of these R&D expenditures were outside of defense. We have been unable to find measures of private R&D spending over this period, but it is almost certain that the sum of private and public R&D spending increased significantly during the war. In particular, the total number of scientists and engineers employed in the manufacturing sector almost doubled between 1940 and 1946. (See Mowery and Rosenberg 2000.)

⁶The appendix describes some of the bond market regulations in detail.

⁷Households held only 5 percent of the federal debt at this time. The appendix discusses these regulations, summarizes the portfolio composition of households versus regulated intermediaries, and shows that the great majority of debt was held by these regulation institutions.

⁸We include the possibility of transfers because it will let us examine how changing the quantity of debt issued by the government (by allowing a fraction of expenditures to be financed with lump-sum taxes) affects the results.

⁹Alternatives to this specification include one used by Ohanian (1997) in which some families were hit by the draft and others were not, and Mulligan (1998), who considers only hours and subtracts draftee hours from the household's total time endowment. Ohanian preserves the representative agent assumption by assuming separable utility between consumption and leisure, and by assuming that labor income for those in the military and private workers was the same. Mulligan's formulation is also a representative agent specification, but it does not allow one to distinguish

between the labor force and hours per worker.

¹⁰We do not include foreign net investment with private investment since the labor input in (6) includes only domestic labor. Instead, we add it with military investment. This choice has little impact on the results since the current account is relatively small in this period.

¹¹McGrattan and Ohanian (2008) evaluate the sensitivity of the results to changes in the starting date of the analysis by beginning in 1941 rather than 1939. The results are very close.

¹²Some readers may be interested in trying to understand the results by decomposing the impact of the shocks into wealth effects and substitution effects, as in Barro (1981) and Hall (1980). This is complicated because the shocks are realized over different points in time, which generates sequences of wealth, intratemporal substitution, and intertemporal effects and intra- and intertemporal substitution effects that are not easy to understand. Our alternative approach of evaluating the contribution of each shock provides a much easier way of understanding the factors that are driving these results.

¹³See the appendix for more details.

¹⁴To construct an after-tax wage, Mulligan uses the Barro-Sahasakul (1986) tax rate that mixes tax rates on labor and capital income. In McGrattan and Ohanian (2008), we rerun our numerical experiments using this tax rate for τ_{lt} and compare it to our benchmark results.

¹⁵There are some measurement differences between our analysis and those of Mulligan (1998) and Rotemberg and Woodford (1992), though they are not central to our findings. Our measure of the price deflator excludes military compensation, which is appropriate for our model. Mulligan deflates the wage using the CPI. Rotemberg and Woodford deflate World War II wages using the GNP deflator, which includes military compensation. For the manufacturing wage, Mulligan uses the wage from 25 industries, whereas we use total compensation for all manufacturing. Finally, we construct an after-tax wage using Joines' (1981) estimate of the labor tax rate. Mulligan uses the Barro-Sahasakul (1986) tax rate. Our main results are robust to these measurement differences.

¹⁶All of the wage measures exclude estimates of wages of farm proprietors because, in general, it is hard to estimate the fraction of proprietor's income that is labor income and, more specifically, because the relative price of farm output nearly doubled during World War II. Accounting for this enormous relative price change is beyond the scope of our one-sector model. It should be noted that the wages of farm employees are included in our full-time equivalent wage measure.

¹⁷Some authors, including Mulligan (1998) and Rotemberg and Woodford (1992), have used average hourly earnings in 25 manufacturing industries as a wage measure for the manufacturing sector. This series has two drawbacks. One is that it does not include non-wage compensation, which was significant in World War II. Also, it does not cover the entire manufacturing sector.

References

- Alchian, Armen. 1963. Reliability of progress curves in airframe production. *Econometrica* 31(4): 679–693.
- Atelson, James B. 1998. *Labor and the wartime state: Labor relations and law during World War II*. Urbana: University of Illinois Press.
- Barro, Robert J. 1981. Output effects of government purchases. *Journal of Political Economy* 89(6): 1086–1121.
- Barro, Robert J., and Chaipat Sahasakul. 1986. Average marginal tax rates from social security and the individual income tax. *Journal of Business* 59(4): 555–566.
- Baxter, Marianne, and Robert G. King. 1993. Fiscal policy in general equilibrium. *American Economic Review* 83(3): 315–334.
- Blanchard, Olivier J., and Roberto Perotti. 2002. An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics* 117(4): 1329–1368.
- Board of Governors of the Federal Reserve. 1944. The wartime expansion of liquid assets. *Federal Reserve Bulletin* 30(10): 953–961.
- Braun, R. Anton, and Ellen R. McGrattan. 1993. The macroeconomics of war and peace. In Stanley Fischer and Olivier J. Blanchard, eds., *NBER Macroeconomics Annual 1993*, pp. 197–247. Cambridge, MA: MIT Press.
- Bullard, Edward. 1975. The effect of World War II on the development of knowledge in the physical sciences. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, A Discussion on the Effects of the Two World Wars on the Organization and Development of Science in the United Kingdom, 342(1631): 519–532.
- Burnside, Craig, Martin Eichenbaum, and Jonas D. M. Fisher. 2004. Fiscal shocks and their consequences. *Journal of Economic Theory* 115(1): 89–117.
- Cole, Harold L., and Lee E. Ohanian. 2002. The Great U.K. Depression: A puzzle and possible resolution. *Review of Economic Dynamics* 5(1): 19–44.
- Cole, Harold L., and Lee E. Ohanian. 2004. New Deal policies and the persistence of the Great Depression: A general equilibrium analysis. *Journal of Political Economy* 112(4): 779–816.
- Davies, D. S., and Judith R. Stammers. 1975. The effect of World War II on industrial science. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, A Discussion on the Effects of the Two World Wars on the Organization and Development of Science in the United Kingdom, 342(1631): 505–518.
- Field, Alexander J. 2003. The most technologically progressive decade of the century. *American Economic Review* 93(4): 1399–1413.

- Goldsmith, Raymond W., ed. 1973. *Institutional investors and corporate stock: A background study*. National Bureau of Economic Research. New York: Columbia University Press.
- Goldsmith, Raymond W., Dorothy S. Brady, and Horst Mendershausen. 1956. *A study of saving in the United States*, Vol. 3. Princeton, NJ: Princeton University Press.
- Gordon, Robert J. 1969. \$45 billion of U.S. private investment has been mislaid. *American Economic Review* 59(3): 221–238.
- Gordon, Robert J. 1970. Communications: \$45 billion of U.S. private investment has been mislaid: Reply. *American Economic Review* 60(5): 940–945.
- Hall, Robert E. 1980. Labor supply and aggregate fluctuations. *Carnegie-Rochester Series on Public Policy* 12(1): 7–33.
- Hansen, Gary D. 1985. Indivisible labor and the business cycle. *Journal of Monetary Economics* 16(3): 309–327.
- Hayashi, Fumio, and Edward C. Prescott. 2002. The 1990s in Japan: A lost decade. *Review of Economic Dynamics* 5(1): 206–235.
- Ibbotson Associates. 2001. *Stocks, bonds, bills, and inflation: Valuation edition*. Chicago: Ibbotson Associates.
- Jaszi, George. 1970. Communications: \$45 billion of U.S. private investment has been mislaid: Comment. *American Economic Review* 60(5): 934–939.
- Joines, Douglas H. 1981. Estimates of effective marginal tax rates on factor incomes. *Journal of Business* 54(2): 191–226.
- Kendrick, John W. 1961. *Productivity trends in the United States*. Princeton, NJ: Princeton University Press.
- Kydland, Finn E., and Edward C. Prescott. 1991. Hours and employment variation in business cycle theory. *Economic Theory* 1(1): 63–81.
- Life Insurance Association of America. 1962. *Life insurance companies as financial institutions*. Englewood Cliffs: Prentice-Hall.
- McGrattan, Ellen R. 1994. A progress report on business cycle models. *Federal Reserve Bank of Minneapolis Quarterly Review* 18(4): 2–16.
- McGrattan, Ellen R. 1996. Solving the stochastic growth model with a finite element method. *Journal of Economic Dynamics and Control* 20(1–3): 19–42.
- McGrattan, Ellen R., and Lee E. Ohanian. 2008. Technical appendix: Does neoclassical theory account for the effects of big fiscal shocks? Evidence from World War II. Staff Report 315, Federal Reserve Bank of Minneapolis.
- McGrattan, Ellen R., and Edward C. Prescott. 2003. Average debt and equity returns: Puzzling? *American Economic Review*, Papers and Proceedings, 93(2): 392–397.
- Mowery, David, and Nathan Rosenberg. 2000. Twentieth century technological change. In Stanley

- L. Engerman and Robert E. Gallman, eds., *The Cambridge economic history of the United States*, Vol. 3. Cambridge, UK: Cambridge University Press, pp. 803–926.
- Mulligan, Casey B. 1998. Pecuniary incentives to work in the United States during World War II. *Journal of Political Economy* 106(5): 1033–1077.
- Ohanian, Lee E. 1997. The macroeconomic effects of war finance in the United States: World War II and the Korean War. *American Economic Review* 87(1): 23–40.
- Rockoff, Hugh. 1984. *Drastic measures: A history of wage and price controls in the United States*. Cambridge, UK: Cambridge University Press.
- Rotemberg, Julio J., and Michael Woodford. 1992. Oligopolistic pricing and the effects of aggregate demand on economy activity. *Journal of Political Economy* 100(6): 1153–1207.
- Siu, Henry E. 2008. The fiscal role of conscription in the U.S. World War II effort. *Journal of Monetary Economics*, 55(6): 1094–1112.
- U.S. Department of Commerce. Bureau of the Census. 1975. *Historical statistics of the United States: Colonial times to 1970*. Bicentennial ed., Part 2. Washington, DC: U.S. Government Printing Office.
- U.S. Department of Commerce. Bureau of Economic Analysis. 1986. *The national income and product accounts of the United States, 1929–82*. Washington, DC: U.S. Government Printing Office.
- U.S. Department of Commerce. Bureau of Economic Analysis. 1987. *Fixed reproducible tangible wealth in the U.S., 1925–85*. Washington, DC: U.S. Government Printing Office.
- U.S. Department of Commerce. Bureau of Economic Analysis. 2002. Fixed assets and consumer durable goods for 1925–2001. *Survey of Current Business* 82(9): 23–37. www.bea.gov.
- Wasson, Robert C., John C. Musgrave, and Claudia Harkins. 1970. Alternative estimates of fixed business capital in the United States, 1925–1968. *Survey of Current Business* 50(4): 18–36. Washington, DC: U.S. Government Printing Office.

TABLE 1
PARAMETER VALUES FOR MODEL SIMULATIONS

Preferences	$\psi = 2.37, \xi = .0, \beta = .985, \bar{l}_d = 50/84$
Technology	$\theta = .38, \delta = .055$
Growth	$\gamma_n = .015, \gamma_z = .02$

TABLE 2
NONMILITARY HOURS DURING WORLD WAR II^a

Year	Model Predictions				U.S. Data
	Government Spending Only	Add Technology Shocks	Add Military Draft	Add Distortionary Taxes	
1939	128	110	111	107	97
1940	130	113	115	109	100
1941	131	122	123	114	106
1942	132	129	127	114	112
1943	131	132	127	111	116
1944	128	132	125	109	113
1945	124	128	122	105	105
Average	129	124	121	110	107

^a All series are relative to average U.S. nonmilitary hours for 1946–1960.

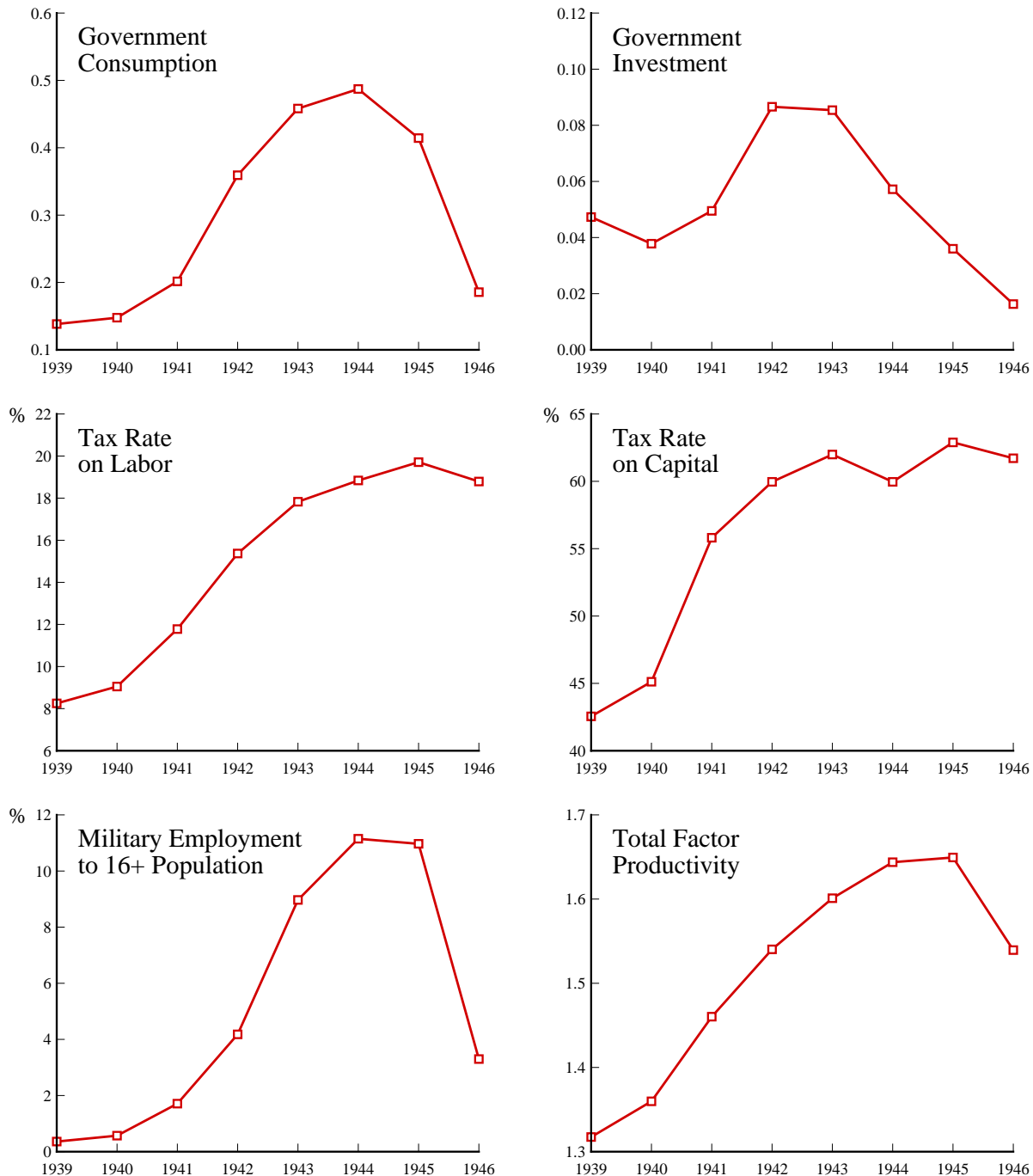
TABLE 3

AFTER-TAX WAGES DURING WORLD WAR II

Year	Model Predictions				U.S. Estimates, Based on			
	Government Spending Only	Add Technology Shocks	Add Military Draft	Add Distortionary Taxes	NIPA Product	NIPA Compensation		
					Civilian	Nonfarm	Manufacturing	
1939	91	96	96	97	101	96	106	111
1940	91	97	97	98	101	96	105	110
1941	89	98	98	100	102	97	102	105
1942	86	98	99	101	101	99	102	104
1943	83	99	100	103	100	101	103	103
1944	82	100	102	103	102	102	104	103
1945	80	100	101	101	102	104	106	103
Average	86	98	99	100	101	99	104	105

^a Model series and U.S. labor productivity are relative to average U.S. labor productivity for 1946-1960; U.S. estimates based on NIPA incomes are relative to own averages for 1946-1960 with the exception of manufacturing, which ends in 1957.

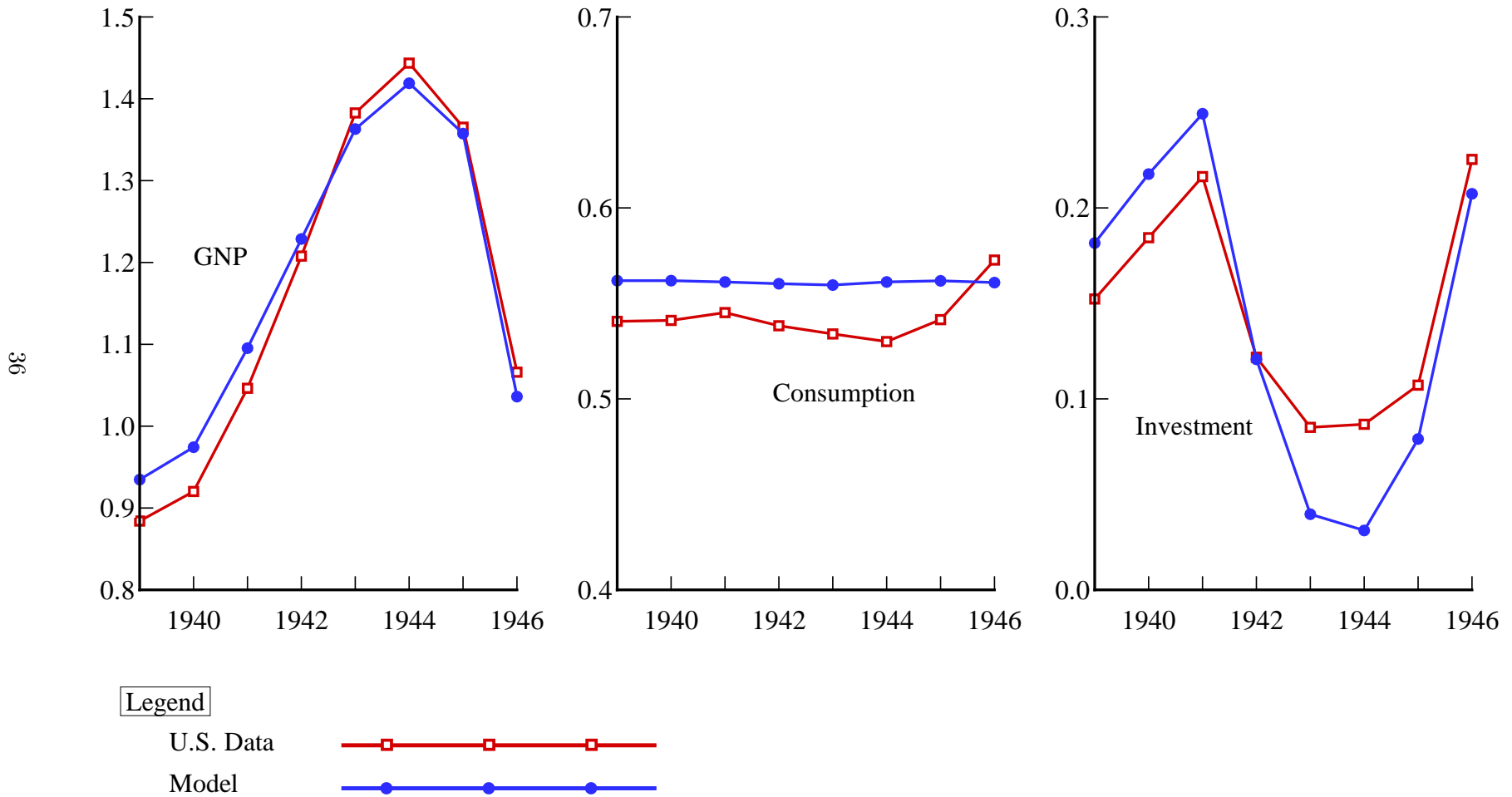
Figure 1. U.S. Government Spending, Tax Rates, Draft, and TFP, 1939–1946



Notes:

- (1) Government spending series are real and detrended by dividing by the population over 16 and by the growth trend in technology (scaled so that the 1946 real detrended level of GNP less military compensation equals 1).
- (2) Total factor productivity is defined to be $Y/(K^\theta L_p^{1-\theta})$, where Y is real detrended GNP less military compensation, K is real detrended nonmilitary capital stock, L_p is nonmilitary hours worked, and $\theta = .38$.

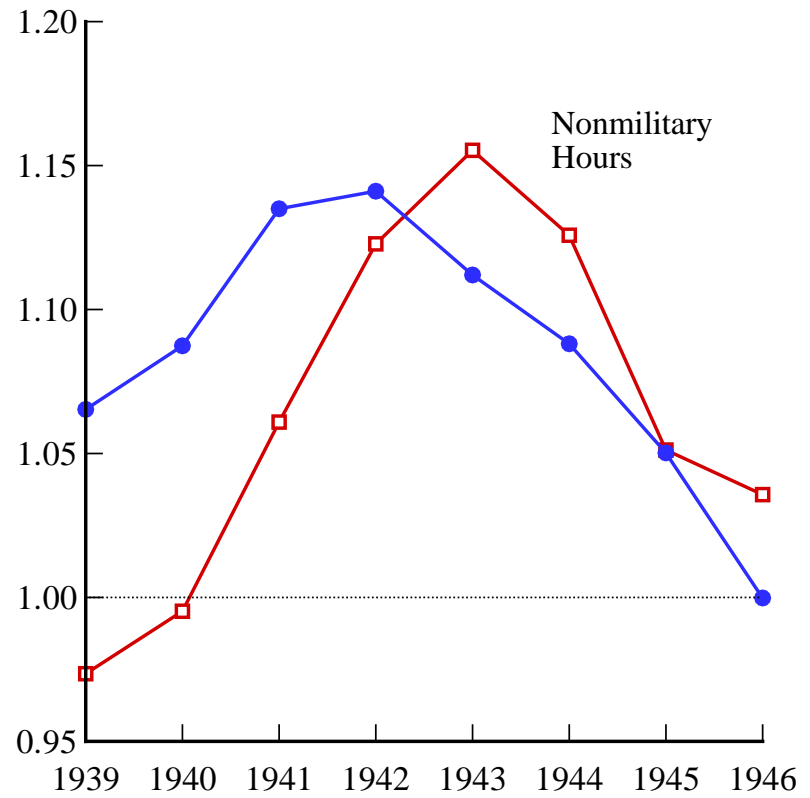
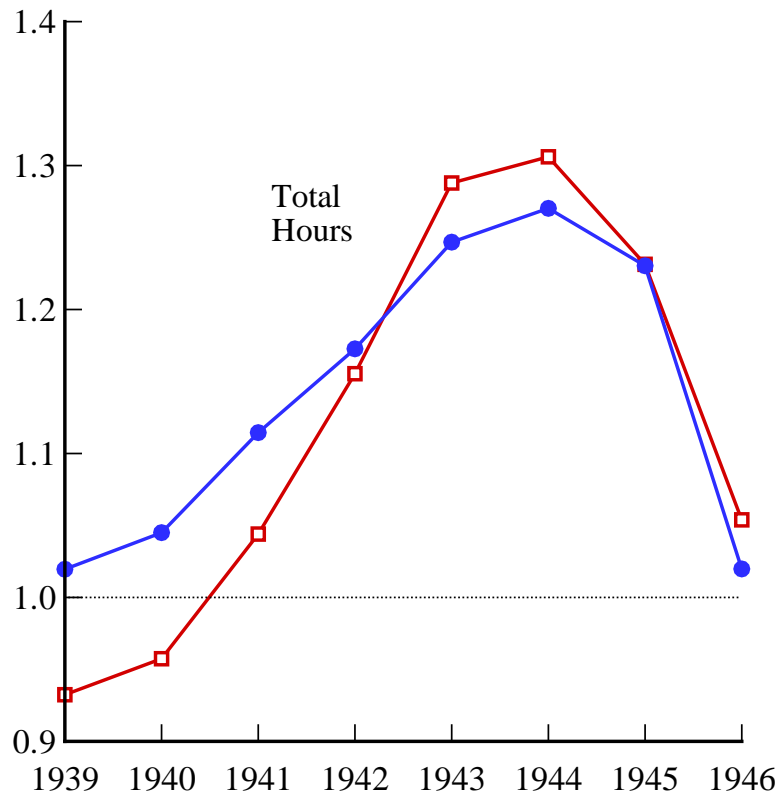
Figure 2. Real Detrended GNP, Private Consumption, and Private Investment, 1939–1946
(Benchmark Deterministic Model)



Note: Data series are divided by the 1946 real detrended level of GNP less military compensation.

Figure 3. Per Capita Total and Nonmilitary Hours of Work, 1939–1946
(Benchmark Deterministic Model)

37

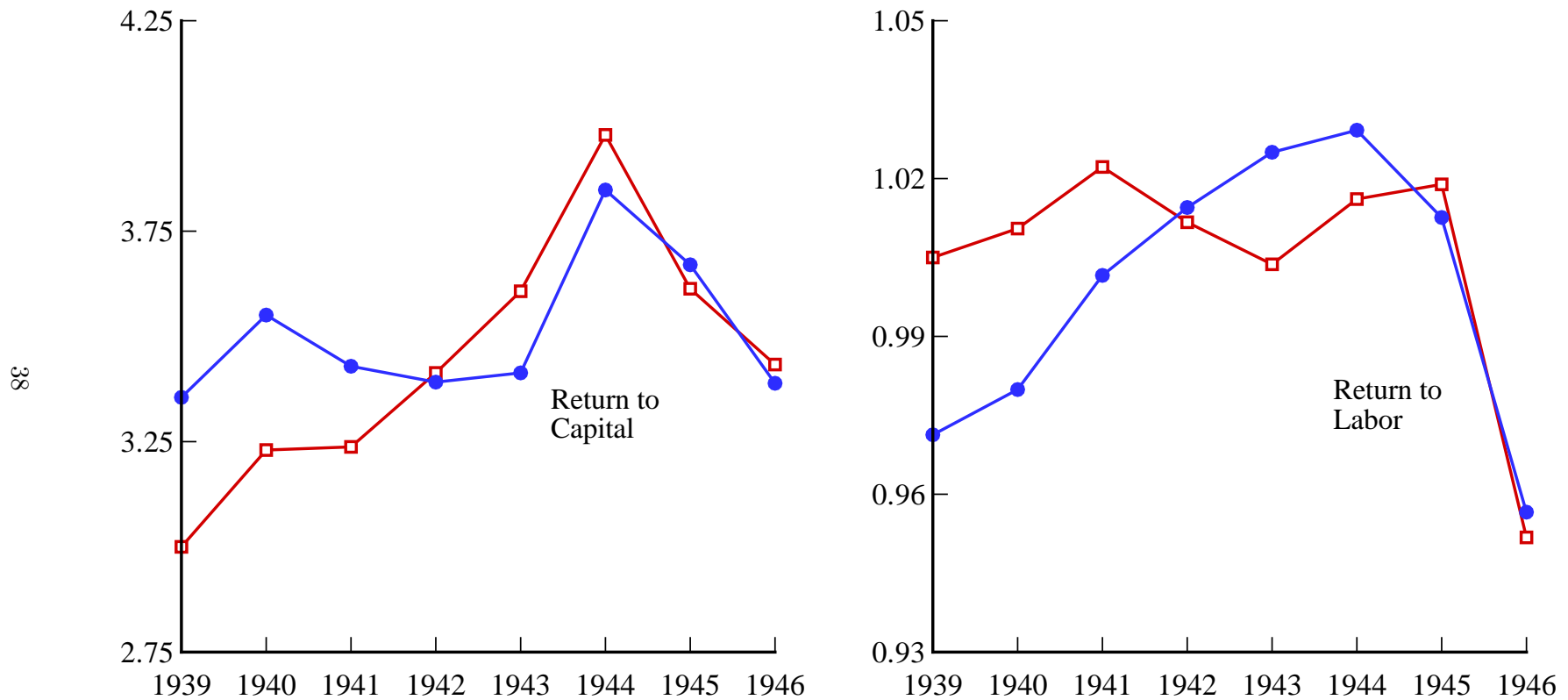


Legend

- U.S. Data —□—
- Model —●—

Note: Hours series are divided by the 1946–1960 U.S. averages.

Figure 4. After-tax Returns to Capital and Nonmilitary Labor, 1939–1946
 (Benchmark Deterministic Model, All Series Constructed Using Marginal Productivities)



Legend

- U.S. Data —□—
- Model —●—

Note: Return to capital is equal to $100(1-\tau_k)(\theta Y/K-\delta)$.

Return to labor is after-tax nonmilitary labor productivity, with hours normalized by the 1946–1960 U.S. average.

Figure 5. U.S. Pre-tax Wage Rates and Prediction of Model with Labor Tax Rate Series Implying Nonmilitary Hours Are Equal in Model and Data, 1939–1946

