Federal Reserve Bank of Minneapolis
Research Department Staff Report 315

Revised March 2007

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, and Julio Rotemberg 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. Between 1941 and 1944, real per capita GNP rose 46 percent. 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 shocks—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 shocks 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 shocks, such as the very large changes in tax rates, in government investment, and in the draft? How does uncertainty about the war affect the model's ability to account for the macroeconomy? 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.

This paper addresses these questions within a quantitative, dynamic general equilibrium framework. We construct a neoclassical growth model tailored to study World War II by including four types of shocks 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 develop a perfect foresight version of the model for heuristic purposes, and we then develop a stochastic version that serves as the primary model for studying the wartime economy. We analyze the stochastic model with all of the shocks to determine the conformity of the model to the data, and we then conduct stochastic experiments that include one shock at a time to isolate the contribution of each shock to the wartime economy. We also use the stochastic model to test whether the expectation of a postwar Depression has an important effect on wartime economic activity. Our stochastic analysis also evaluates how plausible changes in expectations over the exogenous shocks affect the results. In this regard we provide a new Monte Carlo–based approach to choosing probabilities over the states of a model when there is either insufficient information to precisely choose those probabilities, or when the analyst is interested in the sensitivity of quantitative results to changes in probabilities. This new procedure can be used in any model that has a stationary Markovian representation.

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 declines in consumption and investment, and the wartime changes in factor prices. The most surprising finding is that the deviations between the model variables and the actual data during World War II—a period of enormous shocks and many economic regulations and restrictions—are about the same size as those reported in studies that have analyzed the post–World War II period, when the shocks are much smaller and regulations are less important.

Regarding the relative importance of the different wartime shocks, we find that the most important shock by far is the large increase in government spending, followed by productivity change. The very large changes in the draft and in capital and labor income tax rates have much smaller effects. We also find that the results are robust to a wide range of uncertainty about the state variables in the model.

The paper is organized as follows. Section 2 reviews the literature and defines the scope of our analysis. Section 3 summarizes the shocks that we feed into the model. Section 4 summarizes the economic restrictions and regulations adopted during World War II. Section 5 presents the model. Section 6 presents the specification of the exogenous processes and parameter values. Section 7 conducts the quantitative analysis by comparing the model to the data. Section 8 discusses our findings in light of the issues raised in the literature. Section 9 summarizes the sensitivity analysis we conduct. Section 10 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 outcomes to the data. This paper will address these questions, which are summarized below.

There are a number of questions about labor supply and post-tax wages. Mulligan (1998), Baxter and King (1993), and Burnside, Eichenbaum, and Fisher (2004) all question the ability of the model to account for labor supply and after-tax factor price changes during World War II and/or during other episodes of higher military spending and higher taxes. All of these economists are broadly motivated by the fact that two opposing forces impact labor 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 will have a difficult time accounting 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 may require "patriotism," modeled as preference shifts, to jointly account for wartime labor supply and factor prices. Baxter and King (1993) 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) focus on the impact of post–World War II fiscal shocks on labor supply and after-tax wages. They argue that the standard neoclassical model (the one we use here) cannot quantitatively account for the impact of post–World War II fiscal shocks on real wages and labor supply, and that habit formation and investment adjustment costs are required. We will therefore compare labor supply, wages, and capital returns from the model to their data counterparts to address these questions.

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.

There are questions about changes about 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.

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 shocks, then feed the shocks into a parameterized model and 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 1941–1946.

We focus our analysis on World War II because it is the most important and striking 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 also very interesting to study but are beyond the scope of this paper.[4]

3. World War II Shocks

This section summarizes the shocks that we include in the model. The three types of government policy (fiscal) shocks: government consumption and investment, income tax rates, and the draft. We call these shocks the *fiscal state* of the economy. In addition to these fiscal shocks, we also include productivity shocks. Figure 1 displays all six shocks.

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 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 almost doubled, rising from about 10 percent to about 19 percent. Capital tax rates rose about 50 percent, from about 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 about 1 percent before the war to more than 12 percent at the peak of the war. 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 shocks, 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 2 percent per year. Detrended TFP rises 13 percent between 1941 and 1945. 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, plausibly 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. Sir Edward 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 about 18 percent, with relatively high gains in several of the sectors noted above that were affected by R & D, including chemicals and allied products (up 49 percent), electrical machinery (up 26 percent), instruments and miscellaneous (up 25 percent), and primary metals (up

41 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 about 2.5 percent between 1946 and 1948, and rose by 8.3 percent between 1948 and 1950. 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 82 percent, 61 percent, 45 percent, and 35 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 advance. 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.

Figure 1 also shows that productivity declined after the war was over. This decline is not surprising, 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 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 productivity growth may have been 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 in our sensitivity analyses a version of our model with variable capacity.

4. Wartime Economic Restrictions and Regulations

A number of economic restrictions and regulations were adopted during World War II. 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 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. Appendix A describes some of the bond market regulations in detail.

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 simple environment in the standard growth model. 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 (households held only 5 percent of the federal debt at this time. Appendix A 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. Including these regulations would clearly add significantly to the complexity of the model.

Because these additional factors would complicate the analysis considerably, and because we do not know what their impact would be, we abstract from them in our initial analysis and assess how well a neoclassical model without (*i*) coupon rationing and price and wage controls, without (*ii*) constrained investment in certain sub-sectors of the economy, and without (*iii*) control of the pricing of government debt can account for the U.S. data during World War II. This approach is a natural first step, because if any of these abstractions are quantitatively important, then the model will fail significantly along one or more dimensions, and this would indicate how the model should be modified.

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, and government payments to military personnel, it includes taxation of capital and labor income, and it includes 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}). 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 marginal choices made by civilians. This is important because there is considerable uncertainty over this marginal disutility of labor for the draftees, because there is no standard estimate of 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.[6]

The family optimally chooses consumption of both types, which we denote by c_c and c_d, respectively. The family also chooses private investment in physical capital, i_p, and civilian labor input, l_c, 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

$$\text{(1)} \qquad \max_{\{c_{ct}, c_{dt}, i_{pt}, l_{ct}, b_{t+1}\}} E_0 \sum_{t=0}^{\infty} \beta^t \ \{(1 - a_t)U(c_{ct}, l_{ct}) + a_t U(c_{dt}, \bar{l}_d)\} \, N_t$$

subject to

$$\text{(2)} \qquad (1 - a_t)c_{ct} + a_t c_{dt} + i_{pt} + b_{t+1}$$
$$= (1 - \tau_{kt})r_{pt}k_{pt} + (1 - \tau_{lt})w_t(1 - a_t)l_{ct} + \tau_{kt}\delta k_{pt} + R_t b_t + a_t w_t \bar{l}_d + T_t$$

$$\text{(3)} \qquad k_{pt+1} = [(1 - \delta)k_{pt} + i_{pt}]/(1 + \gamma_n)$$

$$\text{(4)} \qquad N_t = (1 + \gamma_n)^t$$

$$\text{(5)} \qquad i_{pt} \geq 0,$$

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, b_{t+1} is new government debt holdings, and T_t are government transfers.[7] All quantities are in per capita terms; the constant growth rate of the population 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.

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

$$\text{(6)} \qquad Y_t = z_t F(K_{pt}, K_{gt}, A_t L_{pt}),$$

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, A_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 grows at the constant rate γ_A: $A_t = (1 + \gamma_A)^t$. The term z_t is a productivity shock. We will specify the process for this shock and the others later in this section.

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 $N_t a_t w_t \bar{l}_d$. Total government spending is the sum of the three expenditure items:

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

Government capital evolves according to the following law of motion:

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

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

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

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 2006) shows that the results are not sensitive to alternative specifications of postwar tax rates that retire this wartime debt.

Five of the exogenous variables in the model have already been discussed: conscription (a_t), the tax rate on capital income (τ_{kt}), the tax rate on labor income (τ_{lt}), government consumption (C_{gt}), and government investment (I_{gt}). The sixth exogenous variable is related to the state of the postwar economy, which we denote as D_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 D_t. The state variable s_t is modeled as a Markov chain. This specification is used for both the deterministic and the stochastic versions of the model. If individuals have perfect foresight, the process s_t is degenerate. In the stochastic economies we specify the transition probabilities over s_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)–(3); (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:

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

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

(iv) the resource constraint

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

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 (2006).

6. Parameterization

This section presents the functional forms and parameter values, and the specifications of the states for the exogenous processes and the associated transition probabilities.

A. Functional Forms and Parameter Values

Functional forms and parameter values are identical in both the deterministic and stochastic economies. Table 1 summarizes the values of all parameters, which we discuss in detail below. Preferences are given by

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

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. We later evaluate the robustness of our results by choosing an alternative 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, but recall from the previous discussion of this formulation that this value does not impact any private choices.

Given that there is both government and private investment, the production technology is

$$(14) \quad zF(k_p, k_g, Al) = z(bk_p^\rho + (1-b)k_g^\rho)^{\frac{\theta}{\rho}}(Al)^{1-\theta}.$$

We assume that government capital and private capital are perfect substitutes ($\rho = 1$). As discussed above, this is a reasonable assumption, as most of this government investment was in government-owned, privately operated plant and equipment. The parameter b governs the relative productivities of government and private capital. Given that the capital are perfect substitutes, we assume that they are equally productive, which implies that $b = 1/2$. We chose $\theta = .38$, 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.

There are two productivity parameters; z is a productivity shifter that fluctuates from its average value, while A grows at a constant rate. The growth rates of trend technological progress (A) and the population (N) are set to their average values over this period: $\gamma_A = 2.0$ percent and $\gamma_n = 1.5$ percent. The stochastic process for z is described below.

B. Specifying the State Vector and Transition Matrix

This section specifies the state vector and the probabilities governing the transitions across those states. We first discuss the realizations of the shocks that comprise the states.

The States

Recall that we will feed in to the model the actual realizations of the exogenous variables. The exogenous variables are the categories of government spending, labor and capital income tax rates, the draft, and productivity. This procedure requires specifying a separate state for each year of the war years (1941–1945), and requires specifying a state for the normal peacetime economy.

To do this, we define a state with 1941 values of the exogenous variables, a state with 1942 values, and so forth through 1945. We also need to construct a peacetime state. For the peacetime state, we use the 1946 values of the exogenous variables.[8] This is also the peacetime state for all but one of the stochastic experiments. The specific values of the exogenous variables used in the experiments are plotted in Figure 1. One noteworthy feature of these data is that the peacetime state has income tax rates that are high relative to those during and before the war.

In one of our stochastic experiments, we will allow for the possibility of a postwar depression. Thus, this one stochastic case has two possible peacetime states: either the 1946 actual values of the variables, or a depression state. We consider the possibility of a postwar depression in this one stochastic experiment because wartime surveys show that individuals as well as professional economists placed some probability on the event that the economy would reenter the Great Depression after the war. In testimony to the U.S. Senate Special Committee on Post-War Economic Policy and Planning in May of 1944, four economists reported that they expected a depression.[9]

Unfortunately, there is not sufficient survey information to completely calibrate the depression state, because the Gallup surveys do not characterize the expected severity or nature of the depression. We therefore choose a depression state such that the level of productivity in the model in this state is equal to the actual average of productivity between 1930 and 1938, which is 14 percent below trend (defined to be the 1929 level times $(1+\gamma_A)^t$). This productivity-driven model of a depression, based on data from the U.S. Great Depression, generates substantially lower employment and output, and is the simplest mechanism for generating a depression in the model.

Transition Matrix for the States

The economy will transit across the states described above. We now discuss the transition probabilities across these states. For both the perfect foresight and stochastic economies we construct a Markovian state vector, S_t, with a Markov transition matrix denoted as Φ, that governs the transitions across these states. The transition matrix in the perfect foresight economy has unity on the first lower diagonal of Φ governing the appropriate year-to-year transitions, i.e., 1941 to 1942, 1942 to 1943, etc., and it has zeroes everywhere else.

To describe the transition matrix for the stochastic economy, first consider the simplest stochastic case in which there is no possibility of a postwar depression. The elements of the state vector can take on six possible values; five values are the actual realizations of each of the state variables from 1941 to 1945, and the sixth value is the peacetime state, which is the 1946 realization for each of the state variables.

The challenge for any stochastic analysis of World War II—or any other episode that deviates substantially from normal periods—is specifying the transition probabilities of this matrix. We first approached this issue by examining survey evidence on expectations about the war and the economy from Gallup and the other major polls. Unfortunately, the surveys do not provide enough information to estimate probabilities over all possible states in the Markov chain, and therefore some alternative procedure needs to be used. Since we are unaware of any established procedures for this type of case, we developed a Monte Carlo procedure in which we generate transition probabilities randomly from a uniform distribution and keep those draws that generate wars that are empirically plausible—those that generate war frequency and war duration similar to actual U.S. wars. We do not place any other restrictions on the transition probabilities other than the war duration and frequency criteria that we detail below. This approach has the additional benefit of evaluating the robustness of the results to changes in expectations and provides a new tool that can be used in any Markovian model that can be rendered stationary. Note that *any* assessment of this period in a stochastic model will need to address the challenging issue of specifying expectations.

The first step in our procedure generates candidate transition probability matrices by drawing random numbers from a uniform [0,1] distribution, ensuring that the probabilities sum to 1. We keep those transition matrices that generate an average duration of a war, the frequency of a war starting, and the average fraction of years in war that are between 70 percent to 130 percent of their historical U.S. averages. These historical averages are 3.7 years, 4.1 percent, and 15.2 percent, respectively.[10] We generated over 18,000 of these matrices that satisfied the frequency and duration criteria. Appendix A describes this procedure in detail.

We compute the equilibrium of the stochastic model for every transition probability matrix that satisfies these frequency and duration criteria. To summarize the results from all of these trials, we will plot for each date upper and lower bounds for each of the endogenous variables, along with the actual data.[11] We will highlight the distribution of the values for these endogenous variables across the trials by shading the region between the minimum and maximum values. Given our focus on reporting bounds on the endogenous variables, we summarize the transition probabilities from all of the draws in a similar fashion by reporting maximum and minimum values for each of the transition probabilities in the matrix Φ. We report these probabilities in Appendix A. The key probabilities are for the five war years, 1941–1945. The transition probabilities along the diagonal that are equal to 1 in the perfect foresight case range from 0 to nearly 1 in the stochastic model. The other transition probabilities for the war years, which are 0 in the perfect foresight case, range from 0 up to about 0.74 in the stochastic model. Thus, the transition probabilities that we use cover most of the possible probabilities in the transition matrix.

We now turn to the stochastic experiment in which we allow for the possibility of a postwar depression. We consider this extension because there is evidence that individuals and economists placed some probability on a postwar depression. We treat this as an extension, rather than as a baseline feature of the stochastic model, because there is relatively little information that we can use to precisely parameterize the severity of this depression.

We use Gallup and Roper surveys of the time to model the expectations of a postwar depression. For example, in a Gallup survey of September 1941 respondents were asked, "Do you think we are likely to have a greater prosperity, or another depression after the present war?" Seventy-eight percent expected another depression. As we report in Table 1, we set the probability of entering the depression state, conditional on leaving the war, at 78 percent in 1941. We used the same or similar questions asked in Gallup and Roper polls for the other years to set the transition probabilities for the postwar state. Appendix A provides further details.

C. The U.S. Data

We compare seven variables from the model to their counterparts in the U.S. data from 1941 to 1946. 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, services, and the service flow from the stock of consumer durables.[12] We compare private investment in the model to U.S. gross private domestic investment plus foreign net investment. 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 technology, which we estimate at 2 percent per year.

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. Because other researchers have questioned the ability of the model to account for nonmilitary hours, we also compare the nonmilitary component in the model and data. The nonmilitary hours series is the same as that used to compute TFP. 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 to its counterpart in U.S. data. In Appendix A, we also discuss several alternative measures of factor returns and how they compare to the model's predictions.

7. Comparing the Model to the Data

We compare the perfect foresight simulation to the data by plotting the actual realizations of output, consumption, investment, labor, and factor productivities between 1941 and 1946 against the model realizations for these variables. We follow the same procedure for the stochastic simulations with the exception that we report the *distribution* of the outcomes for each model by constructing bounds with minimum and maximum values for each endogenous variable at each date, and the interior between these bounds is shaded according to the density of the distribution. Recall that the *realizations* of the state variables are identical for both the deterministic and stochastic economies. Specifically, all the experiments begin with the capital stock equal to its actual value, and the other state variables equal to their 1941 values. Similarly, the realizations for the state variables for 1942–1946 are also equal to their actual values.[13] We begin with the perfect foresight economy because of its simplicity. We calculate the equilibrium in all cases using the finite element method, because the shocks are sufficiently large as to raise questions about the accuracy of a first-order approximation. (See McGrattan 1996.)

A. Results from the Perfect Foresight Model

Figures 2–4 show time series for the model and data between 1941 and 1946.[14] 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 captures much of the timing of these changes as well.

The first plot in Figure 2 is real detrended GNP in both the model and the data. Real GNP rises about 40 percent through the war, with a large decline occurring between 1945 and 1946. The second plot shows private consumption in both the model and the data. Model consumption shows almost no change, while U.S. consumption shows a decline of about 3 percent between 1941 and 1944. (The stochastic results will show that the very flat consumption pattern in this model is a consequence of the perfect foresight assumption, and we will also see that the deviation between the model and the data is smaller in the stochastic case.)

Most of the change in GNP is not due to private consumption. Rather, it is due to the large changes in government spending and in private investment. 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 that.

Figure 3 has two plots, one for total hours and one for nonmilitary hours. Both the data and the model series are divided by the 1946–1960 U.S. average fraction of discretionary time at work. Thus, the figure shows hours of work relative to a postwar trend. In this perfect foresight simulation, predicted hours are high at the start of the war—higher than observed—because higher labor tax rates are perfectly anticipated. But by 1944, total hours in both the model and the data are close to 30 percent above trend. A significant drop in total hours in both the model and the data occurs between 1945 and 1946 with the large reduction in military employment.

A comparison of the two plots in Figure 3 reveals an interesting fact: despite the huge increase in military employment during World War II, nonmilitary hours in this period rose significantly above trend as well. Specifically, U.S. nonmilitary hours rose to 15 percent above trend. This is also the case for the model. Despite the large increase in labor tax rates, the model hours are above trend during the war because the wartime tax revenues are significantly below the wartime government expenditures—and thus households anticipate high taxes after the war to pay off the government debt. The main difference between the model and the data is the timing of the increase: nonmilitary hours in the model peak in 1941, while in the U.S. data, hours peak in 1943. We will see later that some of this difference in timing is due to the perfect foresight assumption, and some is due to the fact that we have abstracted from other wartime economic features, such as wartime changes in New Deal labor policies.

Figure 4 shows measures of factor returns. The first plot shows the time series for the after-tax marginal product of capital given by $100(1-\tau_k)(\theta Y/K-\delta)$, where Y is nonmilitary output and $K = K_p + K_g$ is capital used to produce nonmilitary output. We multiplied by the capital share θ times one minus the tax rate $1-\tau_k$ and subtracted the depreciation rate in order to put it in the standard units of an asset return. The model predicts a return that is roughly 3.5 percent for the first half of the war, rising to almost 4 percent in 1944. The actual return is similar, varying between 3.25 percent and 4 percent. The second plot in Figure 4 is detrended nonmilitary labor productivity given by $Y_t/[L_{pt}(1+\gamma_A)^t]$. As before, nonmilitary hours L_p are normalized; they are divided by the 1946–1960 average fraction of discretionary time in work. During the war, both the model and the data show a rise in labor productivity relative to the trend in TFP.

This perfect foresight experiment is a useful heuristic analysis because it shows the response of the model endogenous variables when households have perfect foresight. With these results in mind, we now turn to the stochastic experiments.

B. Results from the Stochastic Model

Figures 5–7 show the variables from the stochastic model with all the shocks. This is our main experiment, which we will subsequently refer to as the *benchmark experiment*. Each graph presents the U.S. data and a shaded region representing the model predictions for all probability matrices satisfying the frequency and duration criteria described earlier. The shading indicates the relative mass of the realizations from the experiments, which are approximately normally distributed. Thus, the darker shading toward the midpoint between the bounds indicates relatively more mass, and the lighter shading closer to the bounds indicates less mass. The maximum value for any variable at any date is the highest value for that variable realized out of the 18,000 simulations. Similarly, the minimum value for any variable at any date is the lowest value for that variable realized out of the 18,000 simulations.

The results have several noteworthy features. One is that the changes in output, consumption, investment, labor, and the marginal products of labor and capital are similar between the benchmark model and the data; the actual data typically lie between the upper and lower bounds of the stochastic model, and in most cases there are small differences between the highest and lowest model prediction in any year. For example, there is at most a 10 percent difference between the upper and lower bounds for nonmilitary hours during the war. The fact that the marginal products of labor and capital in the model are similar to those in the data is particularly interesting given the statutory government regulations in the labor and capital markets. Our findings thus suggest that wage controls during the war did not have a significant effect on aggregate labor productivity. Similarly, the findings suggest that the interest rate controls on government bonds did not have a significant effect on aggregate capital productivity.

Another noteworthy feature of the results is that the general patterns of the variables in the stochastic model are similar to those in the perfect foresight model. The largest difference between the perfect foresight and the stochastic case is that consumption declines in the stochastic case, compared to the very flat pattern in the deterministic case, and that on average, nonmilitary hours in the stochastic case do not rise as much early in the war. Otherwise, the differences between the perfect foresight case and the stochastic case are quite small. This similarity is surprising, given that the probabilities for the Markov transition matrix typically differed substantially from those in the (degenerate) transition matrix for the perfect foresight case. An interesting implication of the similarity between the perfect foresight model and the stochastic model is that the results are not very sensitive to variations in expectations, provided that the expectations are consistent with historical war episodes.

C. Assessing the Relative Success of the Model

The most surprising finding is that the deviations between the model and the data during World War II are about the same size as, and in some cases are smaller than, deviations in real business cycle (RBC) models during peacetime periods when the shocks are much smaller and when regulations and restrictions, such as price and wage controls, rationing, and the requirement that the Federal Reserve purchase federal debt and allocate it to regulated institutions, were either nonoperative or much less important. In other words, a reasonable expectation is that the large wartime shocks combined with policies that restricted the normal functioning of markets would have produced relatively much larger deviations.

We further quantify the deviations between the model and the data by reporting the maximum deviation between model and data, and the root mean square deviation. The maximum deviation between actual GNP and the mean of the stochastic model predictions is 4 percent, and the root mean squared deviation of output in this period is 2.6 percent. For investment, the largest deviation—which is roughly 30 percent—occurs in 1945 when both actual and predicted investment are low, and the root mean squared deviation of investment is 2.1 percent. (See Figure 2.) For total hours, the largest deviation between model and data is 6.4 percent, and the root mean squared deviation of hours is 4 percent . For labor productivity, the largest deviation is 2.6 percent, and root mean squared deviation is 1.5 percent.

We now compare the size of these model deviations from World War II to model deviations from other studies that compute the equilibrium time path of dynamic stochastic general equilibrium models during periods with much smaller shocks. Hansen and Prescott (1993), Plosser (1989), McGrattan and Prescott (2006), and Cooley and Ohanian (1997) are four papers that conduct equilibrium path analyses during the less volatile postwar period.[15] Hansen and Prescott use an RBC model with multiple productivity shocks to study the 1990–1991 recession and the subsequent recovery. They compute the equilibrium of the model between 1985 and 1992 and compare the variables from the model to the data at each date in this period. Labor deviations are as large as 4 percent, labor productivity deviations is as large as 3 percent, and deviations in the share of output allocated to investment are as large as 40 percent. Plosser uses an RBC model to study the 1954–1985 U.S. economy, and presents plots of growth rates of model variables compared to actual variables at each date. The model deviations in levels for labor are roughly as high as 10 percent for labor and 20 percent for the real wage. McGrattan and Prescott (2006) establish that standard growth theory—which abstracts from unmeasured intangible investments such as R&D—generates strongly counterfactual predictions for hours during the 1990s hours boom. For the standard growth model, the deviation between actual hours worked and predicted hours worked are as large as 14 percent over the period 1990–2003 and averages 6 percent per year. Cooley and Ohanian use a two-sector growth model to study post–World War II U.K. growth. Deviations between output and labor productivity are as high as 10 percent and 15 percent, respectively.

Our approach thus far has been to compare model and data by plotting each endogenous variable from the model to its data counterpart. An alternative approach computes deviations in the first-order conditions, with a particular focus on the deviation in the household's first order condition that governs time allocation. We therefore compare the ratio between the marginal rate of substitution between consumption and leisure and the after-tax real wage in the model to that in the data.[16] Deviations in this condition are thus a combination of deviations in consumption, hours worked, and the after-tax marginal product of labor. The deviation in this condition averages 2.5 percent between 1941 and 1946 and reaches a maximum of 7.5 percent during this period.

The magnitude of this deviation is small when compared to periods of smaller shocks, such as normal business cycle fluctuations, or when compared to other outlier periods, such as the Great Depression. Chari, Kehoe, and McGrattan (2002, 2006), Cole and Ohanian (2002), and Mulligan

(2002) report that this deviation ranges between 50 percent to over 100 percent during the Great Depression. The finding of a small deviation in World War II indicates that the household's time allocation and consumption decisions during World War II are much closer to the predictions of the growth model than other episodes with large shocks.

In summary, the deviations between model variables and actual variables, as well as between transformations of those variables through first-order conditions, are often larger during the more tranquil post–World War II period than during World War II. This is surprising because studying the impact of very large shocks in a simple model, which omits many restrictions and regulations, would tend to generate larger, not smaller, deviations than those reported here. An important implication of this finding is that the government programs omitted in the model, including wage and price controls, rationing of nondurables and of some types of capital (autos and housing), and interest rate control of federal debt, are not quantitatively important for understanding the major macroeconomic variables from the perspective of the growth model.

D. The Relative Importance of Each Shock

We now discuss the key economic forces that are driving the results. The easiest way to understand and illustrate the various forces is to simulate the model economy by comparing the results from the benchmark model (with all the shocks) to the results from the model that omits one of the shocks. This comparison shows the contribution of each shock individually to the wartime economy. Specifically, we compute the equilibrium with one exogenous variable held constant at its 1941 value throughout the war, and all the other shocks taking on their realizations each year as in the benchmark model. We do this for each of the shocks individually.[17]

Figures 8–12 show the results of these experiments. Figure 8 is central to understanding the behavior of the World War II macroeconomy; it shows the equilibrium of the model with government spending held constant at its 1941 level through 1942–1945. The model variables with government spending fixed at its 1941 value look very different from those in the experiment in which government spending shocks are equal to their actual values for 1942–1945. Specifically, GNP in the benchmark model rises by more than 40 percent during the war, whereas GNP with government spending counterfactually fixed at its 1941 value rises only about 10 percent. Moreover, consumption and investment in the model with fixed government spending are systematically much too high, and labor is systematically much too low relative to the benchmark experiment. In other words, the model economy without the wartime government expenditure shocks, but with all the other shocks, bears virtually no resemblance to the actual wartime economy or to the benchmark model.

This experiment shows that government spending shocks are the major factor in accounting for macroeconomic changes during World War II, and it also illustrates the impact of the wealth effect of the war. Specifically, note that consumption without the government expenditure increase (see Figure 8) is about 10 percent higher than consumption in Figure 5, and hours worked in these two experiments are also correspondingly different. This large difference in consumption and leisure, and the fact that interest rates are not too different between the two experiments, indicates that the wealth effect of wartime government spending is central for understanding wartime macroeconomic behavior.

We will next see that other shocks, particularly tax rates and the draft, are playing more modest roles in contributing to World War II macroeconomic changes, despite the fact that these shocks are large. Figure 9 keeps the number of draftees constant throughout the war; relative to the benchmark model, there is a slight difference in hours worked and the marginal product of labor, otherwise the variables are extremely similar. Figure 10 keeps the labor income tax constant throughout the war; relative to the benchmark model, there is higher labor supply, a higher after-tax wage, and a slightly higher after-tax return to capital, while the remainder of the variables are quite similar. Thus, higher labor taxes reduced employment, ceteris paribus. Figure 11 keeps the capital income tax rate constant throughout the war; relative to the benchmark model, the after-tax return to capital is higher, and investment is also somewhat higher; the other variables are quite similar. Figure 12 keeps (detrended) productivity fixed at its 1941 value throughout the war. Relative to the benchmark model, labor, consumption, and investment differ, with labor and investment significantly higher at the start of the war and lower at the end of the war, relative to the benchmark model. Thus, an important contribution of productivity is to affect the timing of the changes in the variables during the war. To summarize, the government spending shock is the main driving force behind the wartime macroeconomy, while productivity also has an important effect. The labor tax rate, the capital tax rate, and the draft have relatively small effects.

Figure 13 shows the variables from the stochastic model, which includes the possibility of a postwar depression that never occurs. In the model with the depression possibility, consumption is slightly lower and labor slightly higher, reflecting precautionary motives associated with the negative depression state. Otherwise, the variables are similar. This finding, along with the findings that the bounds are fairly narrow, and that the stochastic and deterministic results are similar, suggests that the results are robust to alternative specifications of uncertainty over the exogenous variables.

8. Discussion

This section discusses our findings in light of the questions that have been raised about the ability of the neoclassical model to account for the impact of wartime fiscal shocks. As discussed in Section 2, a number of authors have questioned whether the neoclassical model is consistent with the behavior of labor supply and factor prices during wars. For example, Mulligan (1998) argues that factor prices are too low to be consistent with higher labor supply during World War II: "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). We computed the equilibrium in response to these shocks and found that the model is indeed consistent with these data. High government spending is the key reason why the model can simultaneously account for both wartime labor supply and after-tax factor compensation. In particular, recall that Figure 8 shows government spending fixed at its 1941 level. This figure shows that in the absence of high government spending, the model generates labor supply and after-tax compensation that differ considerably from the data; model labor is too low, and the model after-tax wage is too high. Thus, the enormous resource drain of wartime government spending is the key factor that accurately generates high wartime labor supply without higher after-tax compensation in the model. To understand this result, note that at the peak of the war in 1944, government spending was about 80 percent of trend output. This means that if labor *had not increased*, private consumption would have dropped considerably. Given concave utility, households in the model respond to this large resource drain by consuming less physical consumption and less leisure.

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 deviation between the pre-tax marginal product of labor in the perfect foresight model and its analogue for U.S. data between 1941 and 1946 is −.3 percent, with the largest absolute deviation during this period equal to 2.6 percent. (See Figure 4.) The average deviation between the pre-tax marginal product in the perfect foresight model and the pre-tax U.S. nonfarm compensation per hour is −.1 percent, with the largest absolute deviation equal to 1.6 percent. (See Figure A1 in Appendix A.) Similarly, we find small deviations between the stochastic model's predictions and the U.S. data.[18]

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 39 percent, compared to 38 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. The average deviations are smaller for the stochastic simulations.

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. This stands somewhat in contrast to our findings, which indicate that these additional features may not be required. Many possible reasons may explain these differences. One possible reason is that isolating the effects of increased military expenditures is very difficult in periods when the military shocks are relatively small. For example, the Carter-Reagan military buildup, which is one of the episodes analyzed by Burnside, Eichenbaum, and Fisher, is very small compared to World War II, and moreover it coincides with other significant nonmilitary fiscal changes, such as tax reform and deregulation. While Burnside, Eichenbaum, and Fisher confront this difficult issue of identification in their analysis, it is unknown whether their estimates of the fiscal shock are unbiased, which is required for analyzing the impact of the shocks. In contrast, shock identification during World War II is easier, because the military buildup is so large and is so clearly due to the war that it swamps the other factors that could affect military spending. Another factor is differences in methodology; Burnside, Eichenbaum, and Fisher use VAR analysis, while our approach is to compare the time paths of actual variables to model variables. Yet another factor is unmodeled differences in economic conditions between the 1940s and the Carter-Reagan years. For example, the Carter-Reagan years included large oil price shocks and was also a period during which investment-specific technological change was quantitatively important, according to some authors (Greenwood, Hercowitz, and Krusell 1997; Greenwood and Yorukoglu 1997). Understanding the differences between Burnside, Eichenbaum, and Fisher and this analysis based on these and other differences is an interesting topic for future research, but is beyond the scope of this paper.

Finally, we note that the model tends to generate higher labor at the very start of the war compared to the data, but is closer to the actual data after that. One hypothesis for this narrowing deviation is changes in labor policies. Specifically, New Deal policies that raised wages and depressed employment were still operative at the start of the war, but then tended to decline after that. In particular, Cole and Ohanian (2004) describe how the National War Labor Board increasingly reduced real wage demands on the part of labor, particularly in manufacturing, over the course of World War II.

9. Sensitivity Analysis

In this section, we discuss four sets of experiments that we ran to assess the robustness of our results. The first two focus on the robustness of the prediction for higher private labor input during World War II. First, we evaluate how high labor tax rates would have to be such that nonmilitary hours worked would not rise during the war. Second, we evaluate how low the labor supply elasticity would have to be such that, on average, nonmilitary hours worked remained at their trend level. We conduct both of these experiments in the simpler deterministic model. The third set of sensitivity experiments focuses on the robustness of our predictions of the effects of fiscal shocks when we introduce variable capacity utilization in our model. Specifically, we allow variation in the number of employed and the number of hours in the workweek, since both rose significantly during the war. The fourth set of experiments starts the analysis in 1939 to include the period of military buildup prior to 1941. We briefly describe the results for each of these experiments; further details can be found in McGrattan and Ohanian (2006).

Regarding the counterfactual tax rate analysis, we conduct two specific tax experiments. In the first, we find a sequence of labor tax rates for the period 1941–1946 that imply no change in predicted nonmilitary hours. In the second, we repeat the experiment but vary capital tax rates instead of labor tax rates. 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 keep nonmilitary hours worked during the war from rising. In Figure 14, we show the labor tax rates for this counterfactual experiment versus those estimated by Joines for the United States. 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. These rates are significantly higher than the U.S. rates. When we vary capital tax rates, we find that even when we set them to 100 percent during the war, we cannot choke off the rise in predicted nonmilitary hours because the capital tax rates do not have a large enough effect on labor supply.

We turn next to the choice of labor supply elasticity. We find that the labor supply elasticity needs to drop by more than a factor of eight relative to the log utility case used in the benchmark model to keep the average value of nonmilitary hours of work equal to its trend level. The elasticity for the log utility case is 2.75 percent compared to 0.32 percent in this experiment.

The elasticities used in these sensitivity experiments are 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.

The third set of experiments allows for variable capacity as in Kydland and Prescott (1991). This experiment allows for endogenous variation in the Solow residual. We therefore remeasure the Solow residual using a production function that has variable capacity utilization. We find that the predictions for the variable capacity model are very similar to those presented for our benchmark model. Because of this similarity, we report all results for the capacity utilization model separately in McGrattan and Ohanian (2006).

The fourth set of experiments begins the analysis in 1939, to determine if the results are sensitive to abstracting from the military buildup that occurred before 1941. McGrattan and Ohanian (2006) presents these results for both the perfect foresight and stochastic economies and shows that the results are quite similar to the benchmark results.

In summary, our sensitivity analysis shows that the results are fairly robust to changes in tax rates, in the labor supply elasticity, to the inclusion of variable capacity utilization, and to changes in the starting date of the analysis.

10. 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 of this size.

This paper conducted the most comprehensive evaluation of the growth model to date in response to the very large World War II shocks. 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, draft, tax rate, and productivity shocks are typically within a few percentage points of those variables in the data.

The most surprising finding is that the deviations between the model and the data during this period of very large shocks and significant restrictions on the market economy (rationing, price controls, wage controls, interest rate control of federal debt) are about the same size as, or smaller than, the empirical deviations in similar models simulated during the much more stable and less regulated postwar period. This is surprising because a reasonable expectation is that simple models generically will perform poorly when confronted with shocks and policies that are far outside the norm. This consensus view is well summarized by David Romer, who noted that "even in normal times, the best model is just a guide. If something extraordinary happens, like either Russia goes under or the stock market goes down by 20 percent… the model's not going to be a reliable guide."(See Altman 2005.) Another surprising finding is that the results are not particularly sensitive to variations in expectations. Specifically, the model variables lie in a fairly narrow range across changes in expectations. Moreover, the procedure developed here to investigate the impact of alternative expectations can be adopted more broadly in models in which it is challenging to calibrate expectations of the state variables.

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, and that it is a useful tool for use in future macroeconomic research on World War II. 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 for future research is understanding *why* these restrictions did not seem to have important effects on the aggregate economy. Another topic for future research is investigating 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 for future research is understanding the contribution of R&D, the development of operations research and management sciences practices, and other factors to the wartime productivity increase. A fourth topic is comparing and contrasting the World War II experience with other U.S. episodes of large fiscal shocks and perhaps international evidence on large fiscal shocks as well. 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). A fifth topic is understanding how the regulations and restrictions on portfolios of various financial institutions contributed to the Federal Reserve's policy of nominal interest rate control of federal debt during the war.

Appendix A. Alternative Measures of Factor Returns

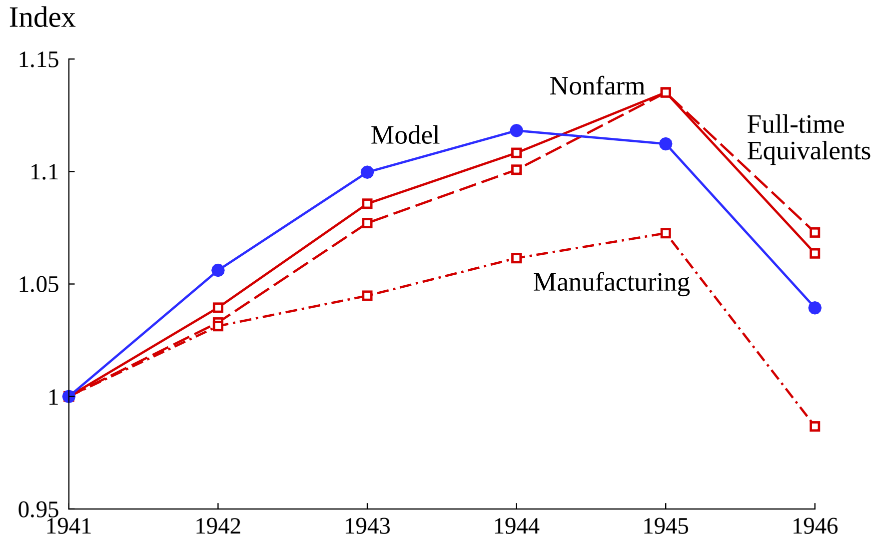
In Section 7 we used marginal products constructed from U.S. data to assess the model's predictions for returns to capital and labor. In theory, we can also use factor incomes per unit of input and, in the case of capital, returns on assets held by capital owners. For labor, we will see that the marginal product from the model accords well with measures of compensation per hour. For capital, we will see that the marginal product on average is consistent with returns based on capital incomes and stocks and with equity returns. We will 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

Figure B1 shows indices of the pre-tax marginal product of labor in our model and three different pre-tax real wage series: nonfarm compensation per hour, full-time equivalent compensation per hour, and manufacturing compensation per hour.[19] All compensation rates are divided first by the deflator for GNP less military compensation and then by the growth trend for technology $(1+\gamma_A)^t$. Along with these compensation rates, we plot labor productivity for the model from Figure 4. Each series is normalized to the value of 100 in 1941.

The two broadest measures of aggregate real compensation per hour—nonfarm and full-time equivalent—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 decline

Figure A1. Compensation Per Hour, Three U.S. Measures and the Benchmark Deterministic Model Prediction, 1941–1946



in manufacturing compensation observed in Figure B1. 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 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 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 a 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 3.57 percent for the perfect foresight 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 comprised 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 companies."
> Life Insurance Companies as Financial Institutions (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 as 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.

Appendix B. Stochastic Simulations

This appendix describes how we generate the transition matrices (Φ) for the Markov chains over the six states used in our stochastic simulations. There are a maximum of seven possible states: a state with 1941 values of exogenous variables, a state with 1942 values of exogenous variables, and so on up to 1945, a postwar state with no depression, and a postwar state with depression. The states for 1942–1945 are called the war states. To keep the exposition as simple as possible, we omit the depression state in the discussion that follows.

To choose the probabilities over these states, we randomly construct transition matrices and keep those that generate (*i*) the frequency of the outbreak of war, (*ii*) the average duration of war, and (*iii*) the fraction of years spent in war, that are between 70 percent to 130 percent of the United States pre–World War II averages for these statistics. (See Table 1.) Thus, households' expectations regarding the frequency and duration of war are equal to the average values from previous U.S. wars. We place no other restrictions on these probabilities so that we can obtain a wide range of expectations and thus check the sensitivity of our results to differences in expectations.

We first generated candidate matrices by drawing each element of Φ independently from the same distribution (and then normalizing columns so they sum to 1). This approach was inefficient because it yielded very few Markov chains that satisfy the above criteria. A much more efficient procedure is to generate probabilities that place relatively more mass on the first off-diagonals of Φ so that the probability of transiting to the next phase of the war is sufficiently high as to obtain more draws that satisfy the above duration and frequency criteria.

In particular, we preset a parameter $\alpha \in [0, .5]$ and then use the following algorithm for $n = 1, \ldots N$, where N is large:

- Initialize elements of the candidate transition matrix, $\Phi_n(i, j) = 0$
- For $j = 1, \ldots 5$,
  - Draw $\zeta \in \text{uniform}[0, 1]$ and set $\Phi_n(j + 1, j) = \alpha + .5\zeta$
  - Draw 5 more uniform random variables on [0,1] for the remaining elements of column j, normalizing them so the probabilities sum to 1

- For $j = 6$,
  - Draw $\zeta \in \text{uniform}[0,1]$ and set $\Phi_n(6,j) = .9 + .1\zeta$
  - Draw 5 more uniform random variables on [0,1] for the remaining elements of column 6, normalizing them so the probabilities sum to 1
- Check to see if candidate Φ_n is within 70 percent to 130 percent of U.S. averages for the above criteria. If so, keep. Continue to next n.

With $\alpha = 0.5$, the algorithm allows off-diagonals to range from 0 to 1, with probabilities over 0.5 more likely. To ensure that we span the probability space, we also repeat the procedure for values of α (discretely) chosen between 0 and 0.5. The probabilities that we construct span most of the relevant probability space in the transition matrix. The following matrix shows the maximum probabilities across all matrices Φ_k that were kept for simulations:

$$\max_k \Phi_k = \begin{bmatrix} .63 & .52 & .65 & .54 & .65 & .04 \\ 1 & .51 & .60 & .58 & .57 & .04 \\ .57 & 1 & .53 & .74 & .53 & .04 \\ .51 & .57 & 1 & .60 & .64 & .04 \\ .54 & .62 & .61 & 1 & .58 & .04 \\ .53 & .51 & .48 & .62 & 1 & .99 \end{bmatrix}.$$

The minimum probabilities are $\min_k \Phi_k(i,j) = .0$ for all elements (i,j) except element (6,6), which has $\min_k \Phi_k(6,6) = .93$. Matrices with low values in element (6,6) violate the criterion governing the duration of the war.

Following the generation of these transition matrices, we then compute the equilibrium for the model economy for each of the accepted transition matrices, feeding in the realizations sequentially from 1941 to 1945, and the postwar realization. Figures 5–13 show the maximum and minimum values for each of the endogenous variables, and the shading in the figure shows the relative mass of the distribution between the minimum and maximum values.

In one experiment, we allow for the possibility of a postwar depression. Participants in a Gallup poll in September of 1941 were asked, "Do you think we are likely to have a greater prosperity, or another depression after the present war?" Seventy-eight percent of respondents said, "Another depression." In June of 1942, Gallup poll participants were asked, "Which do you think the United States will have for the first two or three years after the war—depression or prosperity?" At that time, 43 percent said, "Depression." In June of 1944 and again in May of 1945, a Roper poll asked, "Do you expect we probably will have a widespread depression within 10 years or so after the war is over, or do you think we probably will be able to avoid it?" Fifty-one percent said, "Depression" in 1944 and 44 percent in 1945.[20] We use these wartime survey responses as the model expectations for the postwar state of depression for each year, respectively. (See Table 1.)

Notes

[1]By the *neoclassical model*, we mean the one-sector optimal growth model with a constant returns to scale technology with capital and labor, 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.

[2]A 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.

[3]Rotemberg 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.

[4]Analyzing these smaller episodes is also complicated because exogeneity of the spending, tax, and draft shocks may be harder to justify, and thus would require filtering out the endogenous component of the fiscal changes.

[5]Most 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.)

[6]Alternatives 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 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.

[7]We 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.

[8]We experimented with alternative specifications for the peacetime state by taking an average over multiple years in the post–World War II period. Our findings are robust across these alternatives because the values of the 1946 exogenous variables are very similar to their average

values in the decades following the war.

[9]Testimony was given by Harold Moulton, president of the Brookings Institution; A. F. Hinrichs, acting commissioner of the Bureau of Labor; Matthew Noll of the American Federation of Labor; and Robert Nathan, representing the Commission for Economic Development.

[10]These statistics are calculated from the American Revolutionary War, the War of 1812, the Mexican War, the American Civil War, the Spanish American War, World War I, and World War II.

[11]Specifically, the maximum (minimum) value for a variable at each date is the maximum (minimum) across all realizations for that date.

[12]We subtract indirect business taxes for sales from both GNP and personal consumption expenditures. We impute a service flow for durables equal, in real terms, to 4 percent times the stock of durables.

[13]For our simulations, we ensure that the present value of peacetime tax revenues is sufficient to cover the government debt accumulated during the war.

[14]McGrattan and Ohanian (2006) evaluates the sensitivity of the results to changes in the starting date of the analysis by beginning in 1939 rather than 1941. The results are similar.

[15]Relatively few papers in the RBC literature conduct actual time path analyses in which there is a comparison at various dates between model endogenous variables and actual variables; instead, most of the studies compare some subset of second moments from the model to the data.

[16]See, for example, Parkin (1986), McGrattan (1991), Ingram, Kocherlakota, and Savin (1994), Hall (1997), Cole and Ohanian (2002), Mulligan (2002), Galí, Gertler, and López-Salido (2005), and Chari, Kehoe, and McGrattan (2002, 2006).

[17]Some 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.

[18]There 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. Finally, we construct an after-tax wage using Joines' (1981) estimate of the labor tax rate. Mulligan uses the Barro-Sahasakul (1986) tax rate that mixes tax rates on labor and capital income. Our results are robust to these differences in tax rates and most of the differences in price indices.

[19]All of our 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.

Some 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.

[20]We did not have survey results for 1943 and therefore chose a probability of postwar depression intermediate to our estimates of 1942 and 1944.

References

Alchian, Armen. 1963. Reliability of progress curves in airframe production. *Econometrica* 31(4): 679–693.

Altman, Daniel. 2005. Bernanke's models, and their limits. *New York Times*, October 30.

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: 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 *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.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan. 2002. Accounting for the Great Depression, *American Economic Review*, Papers and Proceedings 92(2): 22–27.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan. 2006. Business cycle accounting. *Econometrica*, forthcoming.

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: 779–816.

Cooley, Thomas F., and Lee E. Ohanian. 1997. Postwar British economic growth and the legacy of Keynes. *Journal of Political Economy* 105(3): 439–472.

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): 505–18.

Field, Alexander J. 2003. The most technologically progressive decade of the century. *American Economic Review* 93(4): 1399–1413.

Galí, Jordi, Mark Gertler, and J. David López-Salido. 2005. Markups, gaps, and the welfare costs of business fluctuations. Economics Working Paper 836, Universitat of Pompeu Fabra.

Goldsmith, Raymond W., ed. 1973. *Institutional investors and corporate stock: A background study.* National Bureau of Economic Research. New York, NY: 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.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell. 1997. Long-Run Implications of Investment-Specific Technological Change, *American Economic Review* 87(3): 342–362.

Greenwood, Jeremy, and Mehmet Yorukoglu. 1997. 1974, *Carnegie-Rochester Conference Series on Public Policy* 46: 49–95.

Hall, Robert E. 1980. Labor supply and aggregate fluctuations. *Carnegie-Rochester Series on Public Policy* 12: 7–33.

Hall, Robert E. 1997. Macroeconomic fluctuations and the allocation of time. *Journal of Labor Economics* 15(1): 223–250.

Hansen, Gary D. 1985. Indivisible labor and the business cycle. *Journal of Monetary Economics* 16: 309–327.

Hansen, Gary D., and Edward C. Prescott. 1993. Did technology shocks cause the 1990-1991 recession? *American Economic Review*, Papers and Proceedings 83(2): 280–286.

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, IL: Ibbotson Associates, Inc.

Ingram, Beth Fisher, Narayana R. Kocherlakota, and N. E. Savin. 1994. Using theory for measurement: An analysis of the cyclical behavior of home production. *Journal of Monetary Economics* 40(3): 435–456.

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 and Edward C. Prescott. 1991. Hours and employment variation in business cycle theory. *Economic Theory* 1(1): 63–81.

Life Insurance Assocation of America. 1962. *Life Insurance Companies as Financial Institutions.* Englewood Cliffs, NJ: Prentice-Hall.

McGrattan, Ellen R. 1991. The macroeconomic effects of distortionary taxation. Discussion paper 37, Federal Reserve Bank of Minneapolis.

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. 2006. 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.

McGrattan, Ellen R., and Edward C. Prescott. 2006. Unmeasured investment and the 1990s U.S. hours boom. Staff Report 369, Federal Reserve Bank of Minneapolis.

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.

Mulligan, Casey B. 2002. A century of labor-leisure distortions. NBER Working paper 8774.

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.

Parkin, Michael. 1986. A method for determining whether parameters in aggregative models are structural. *Carnegie-Rochester Conference Series on Public Policy* 29: 215–252.

Plosser, Charles I. 1989. Understanding real business cycles. *Journal of Economic Perspectives* 3(3): 51–78.

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.

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: 18–36. Washington, DC: U.S. Government Printing Office.

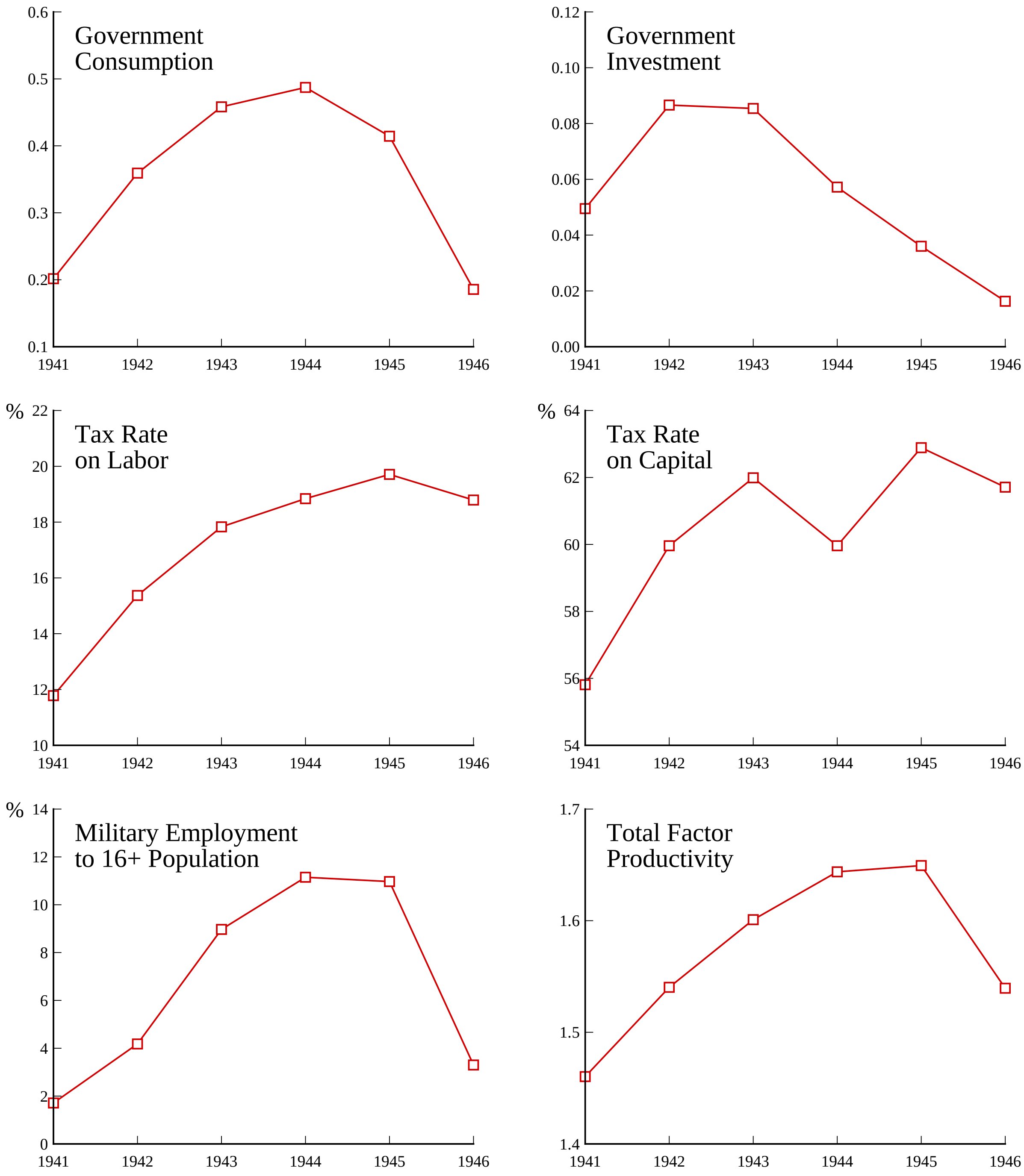
Table 1
Parameter Values for Model Simulations

Preferences	$\psi = 2.37,\ \xi = .0,\ \beta = .985,\ \bar{l}_d = 50/84$
Technology	$b = 1/2,\ \rho = 1,\ \theta = .38,\ \delta = .055$
Growth	$\gamma_n = .015,\ \gamma_z = .02$
Restrictions on Markov Chain†	
Average duration of war	in [2.6, 4.8] years
Fraction of years a war is started	in [2.9, 5.3] percent
Fraction of years in war	in [10.6, 19.8] percent
Probabilities of Postwar Depression†‡	
in 1941	78%
in 1942	43
in 1943	47
in 1944	51
in 1945	44

† Relevant only for stochastic simulations.

‡ The probability is conditional on the war ending.

Figure 1. U.S. Government Spending, Tax Rates, Draft, and TFP, 1941–1946

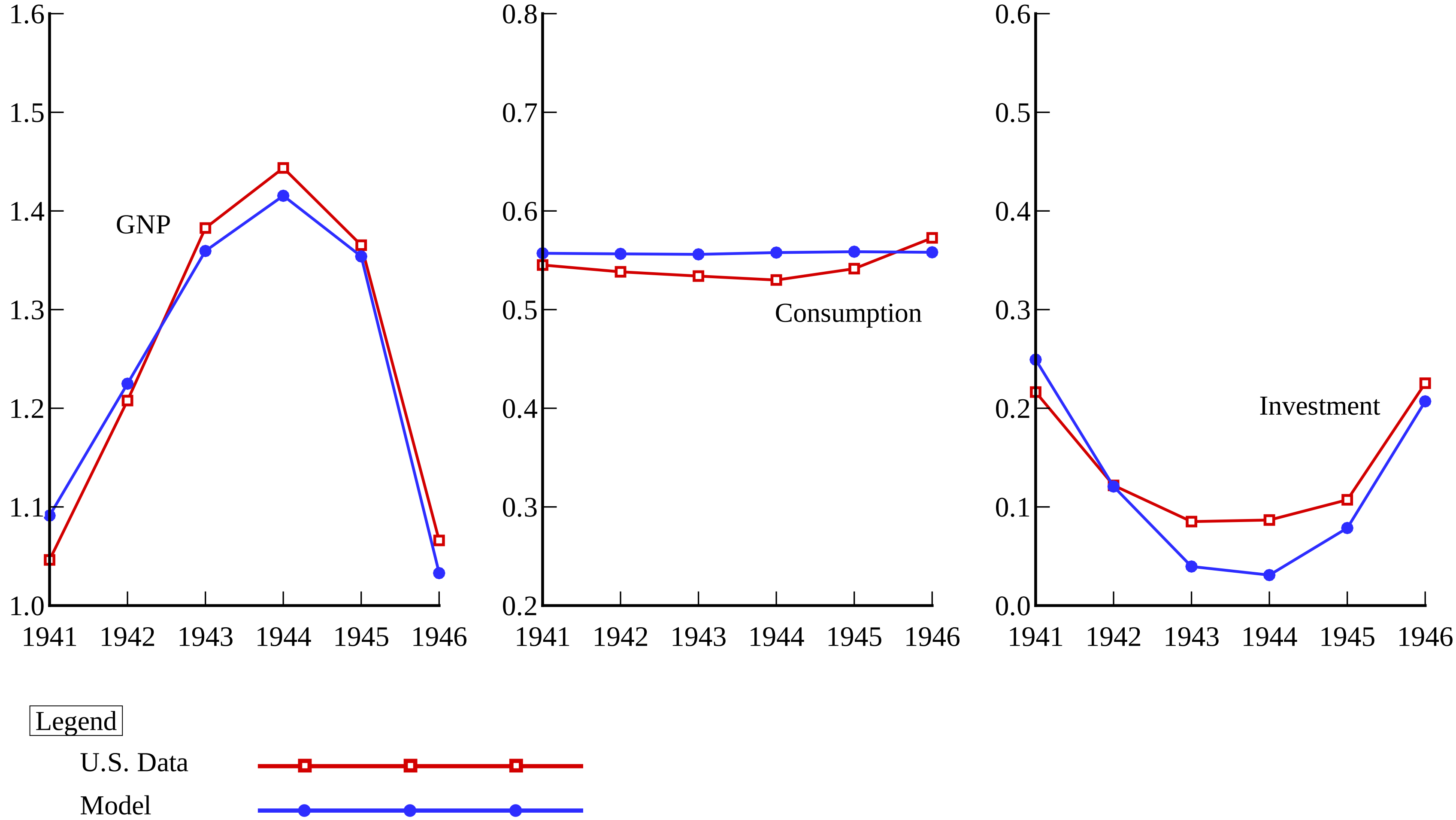


Notes:

(1) Government spending series are real and detrended by dividing by the population over 16 and by the growth trend in technology (scaled so the 1946 real detrended level of GNP less military compensation equals 1).

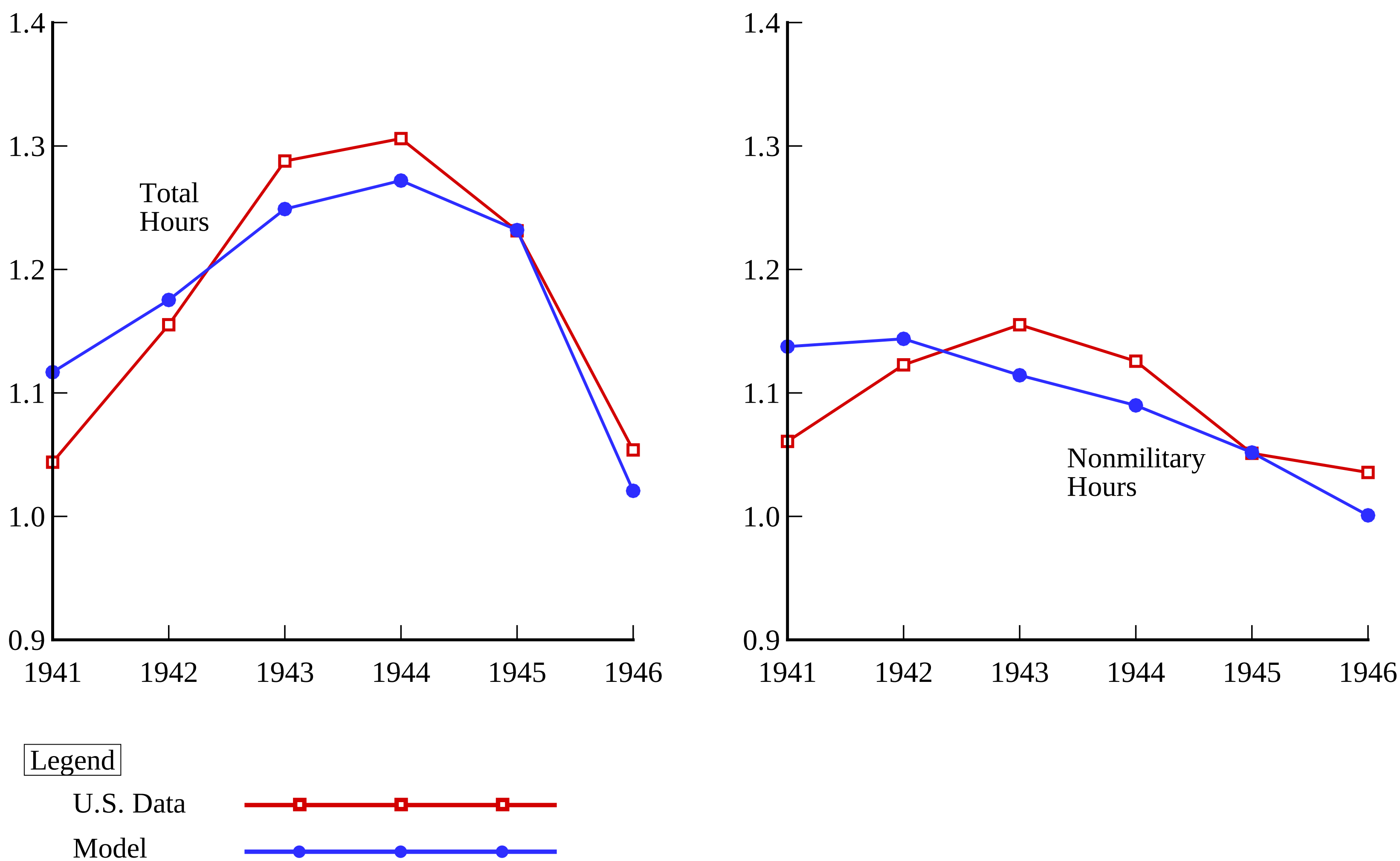
(2) Total factor productivity is defined to be $Y/(K^{\theta}L_p^{1-\theta})$, where Y is real, detrended GNP less military compensation, K is real detrended nonmilitary capital stock, L_p is nonmilitary hours worked, and $\theta = .38$.

Figure 2. Real Detrended GNP, Private Consumption, and Private Investment, 1941–1946
(Benchmark Deterministic Model)



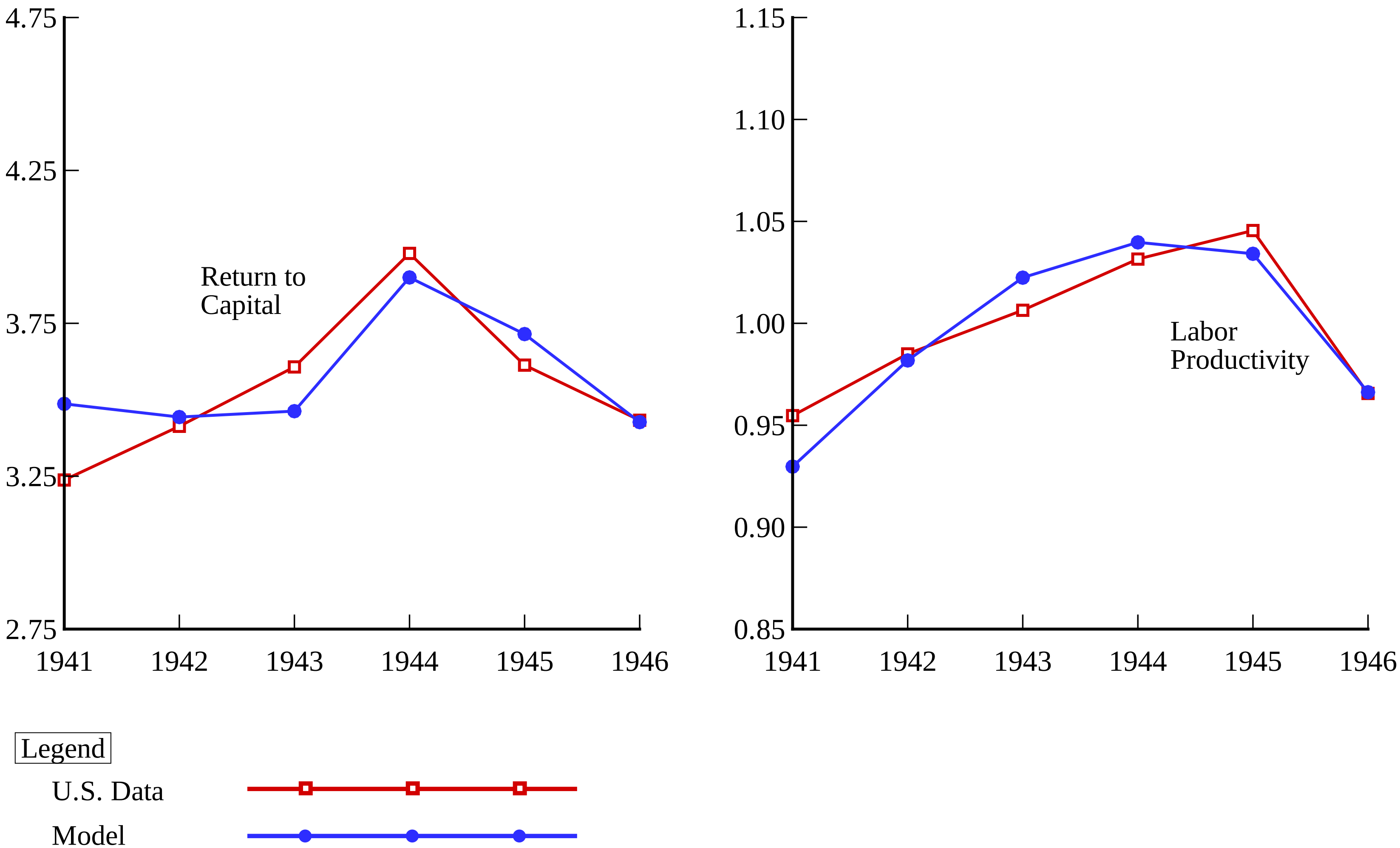
Note: Data series are divided by the 1946 real detrended level of GNP less military compensation.

Figure 3. Per Capita Total and Nonmilitary Hours of Work, 1941-1946
(Benchmark Deterministic Model)



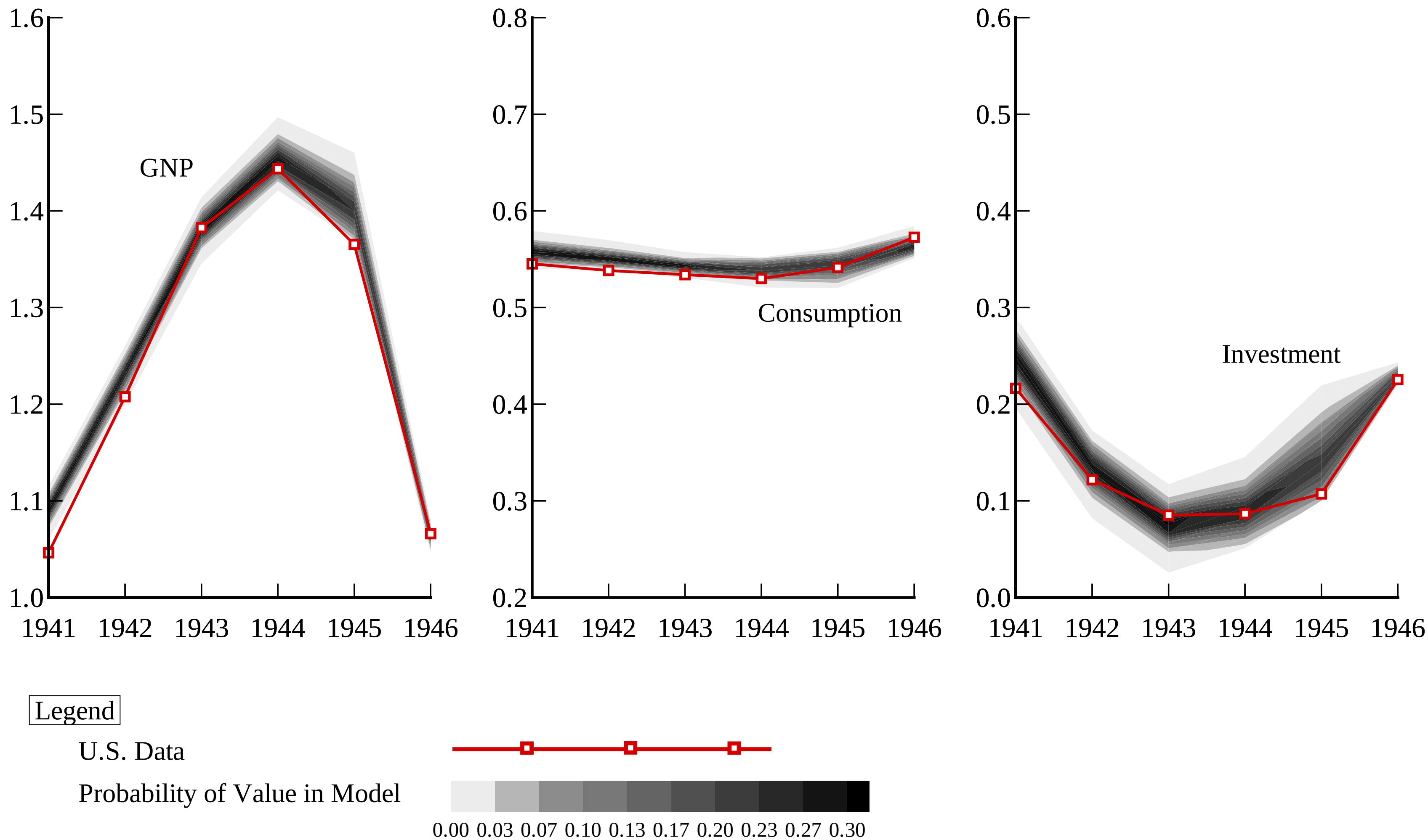
Note: Hours series are divided by the 1946–1960 U.S. averages.

Figure 4. After-tax Return to Capital and Nonmilitary Labor Productivity, 1941–1946
(Benchmark Deterministic Model, All Series Constructed Using Marginal Productivities)



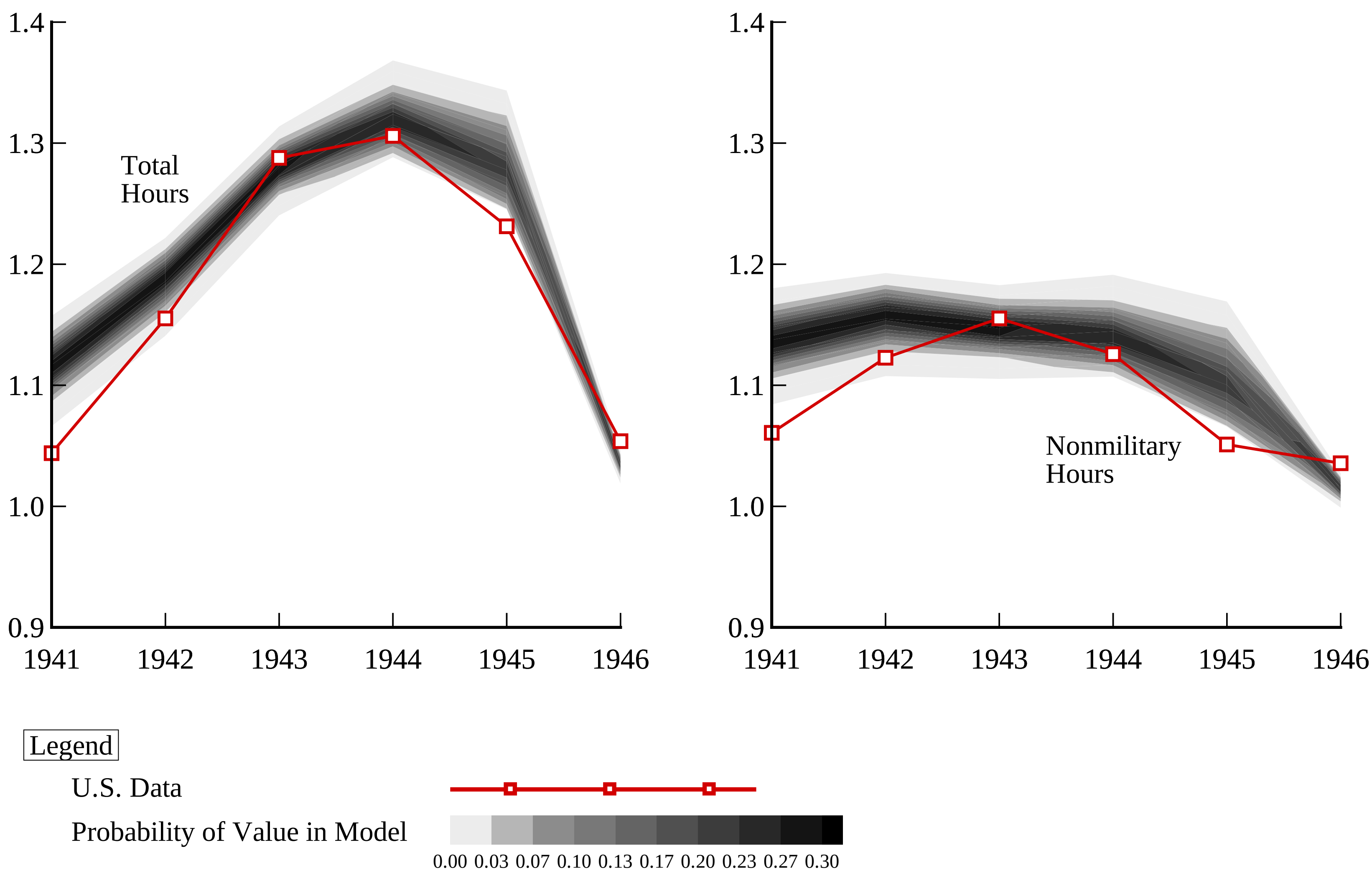
Note: Return to capital is equal to $100(1-\tau_k)(\theta Y/K-\delta)$.
Labor productivity is nonmilitary output divided by hours that are normalized by the 1946–1960 U.S. average.

Figure 5. Real Detrended GNP, Private Consumption, and Private Investment, 1941–1946
(Benchmark Stochastic Model)



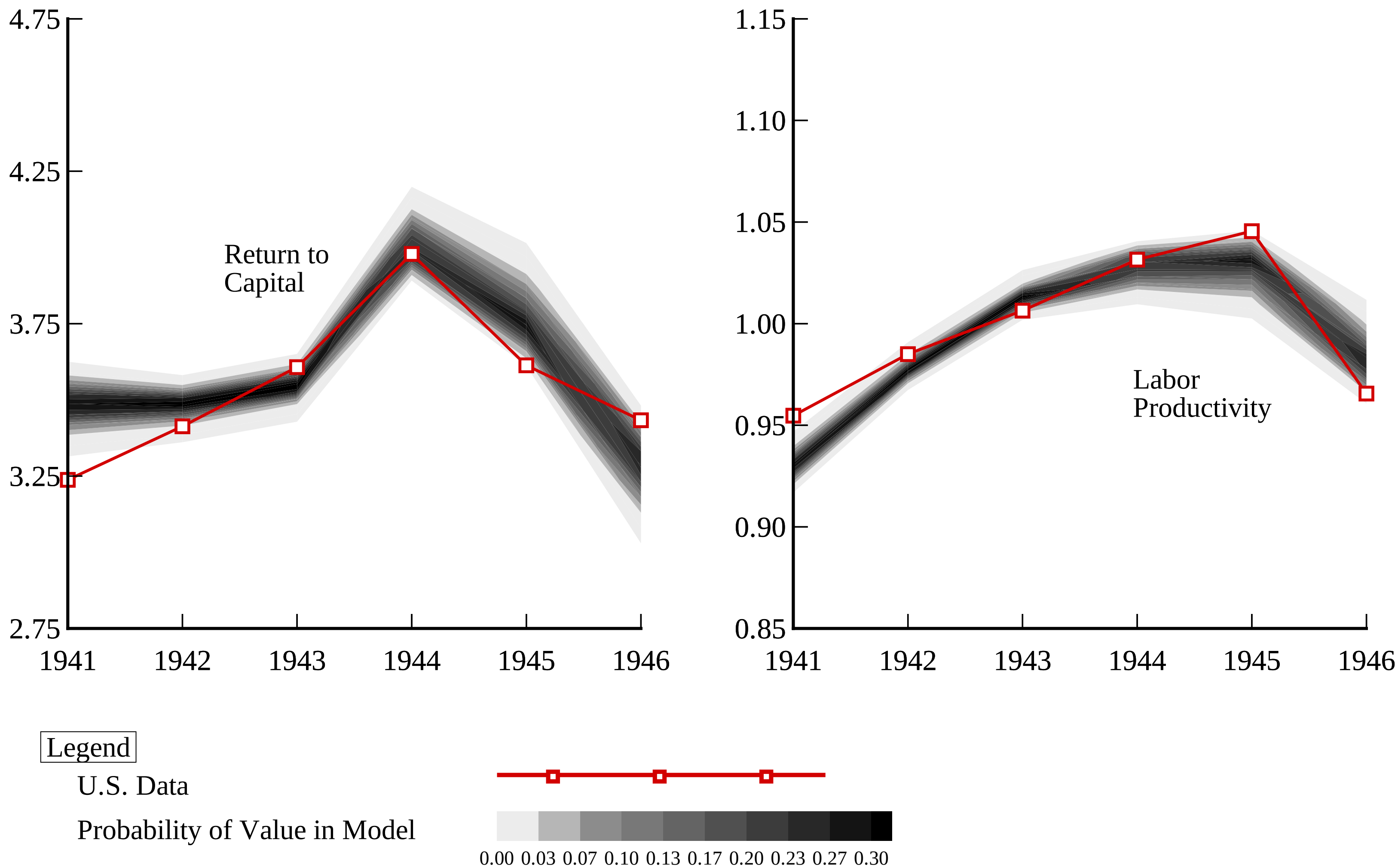
Note: Data series are divided by the 1946 real detrended level of GNP less military compensation.

Figure 6. Per Capita Total and Nonmilitary Hours of Work, 1941–1946
(Benchmark Stochastic Model)



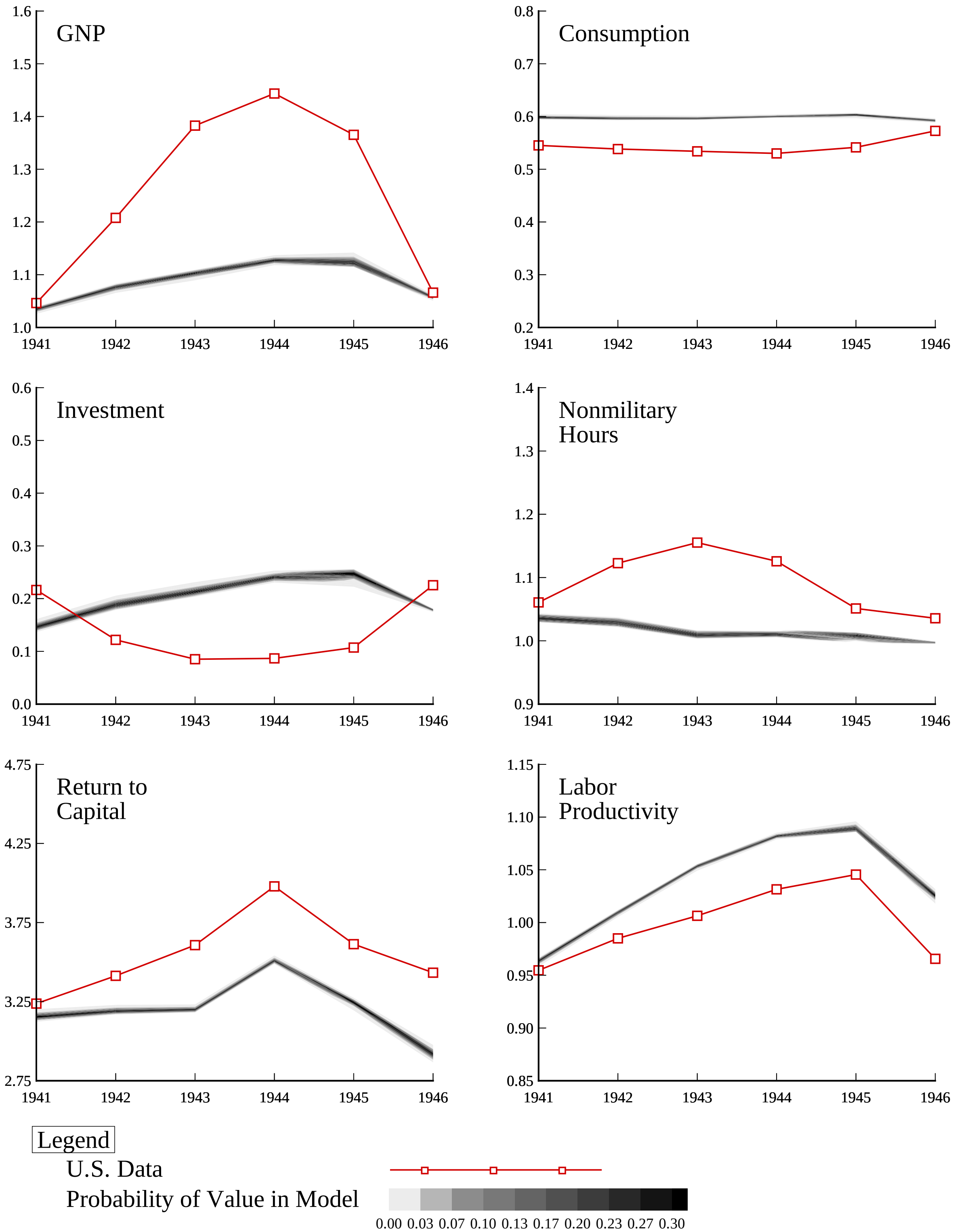
Note: Hours series are divided by the 1946–1960 U.S. averages.

Figure 7. After-tax Return to Capital and Nonmilitary Labor Productivity, 1941–1946
(Benchmark Stochastic Model, All Series Constructed Using Marginal Productivities)



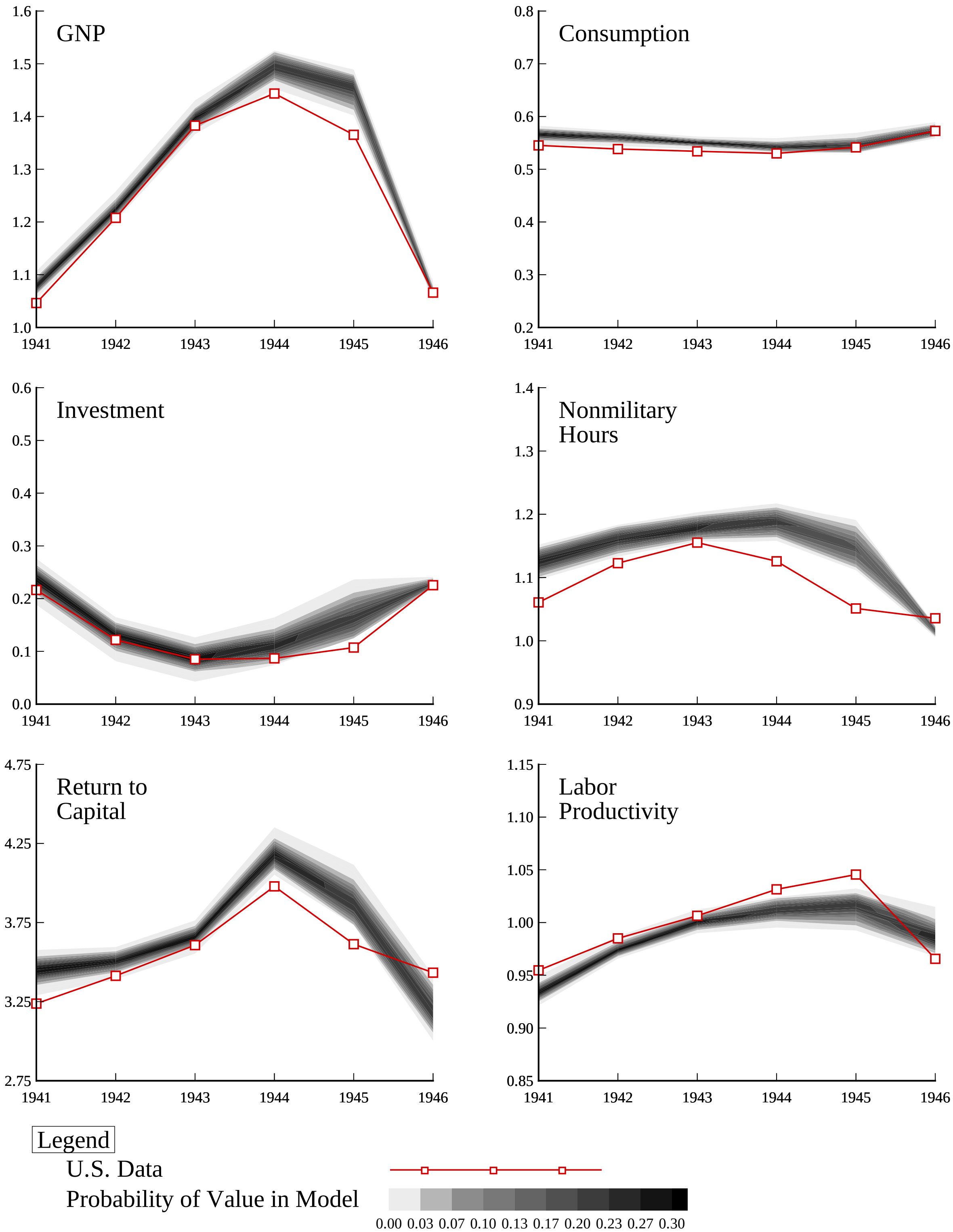
Note: Return to capital is equal to $100(1-\tau_k)(\theta Y/K-\delta)$.
Labor productivity is nonmilitary output divided by hours that are normalized by the 1946–1960 U.S. average.

Figure 8. U.S. Data and Stochastic Model Predictions, 1941–1946
(Model with Government Spending Constant)



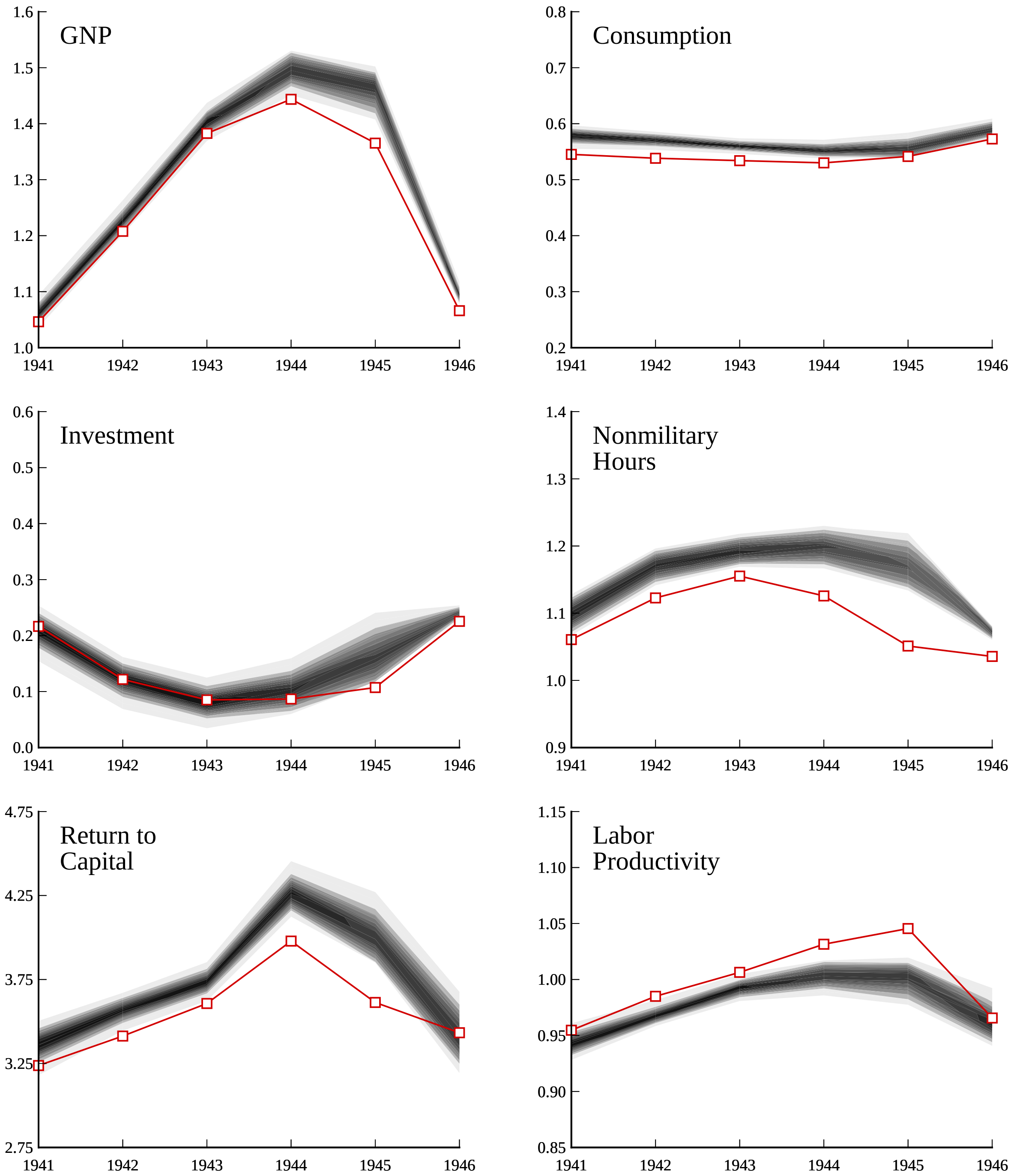
Note: See notes to Figures 5–7 for series normalizations.

Figure 9. U.S. Data and Stochastic Model Predictions, 1941–1946
(Model with Number of Draftees Constant)



Note: See notes to Figures 5–7 for series normalizations.

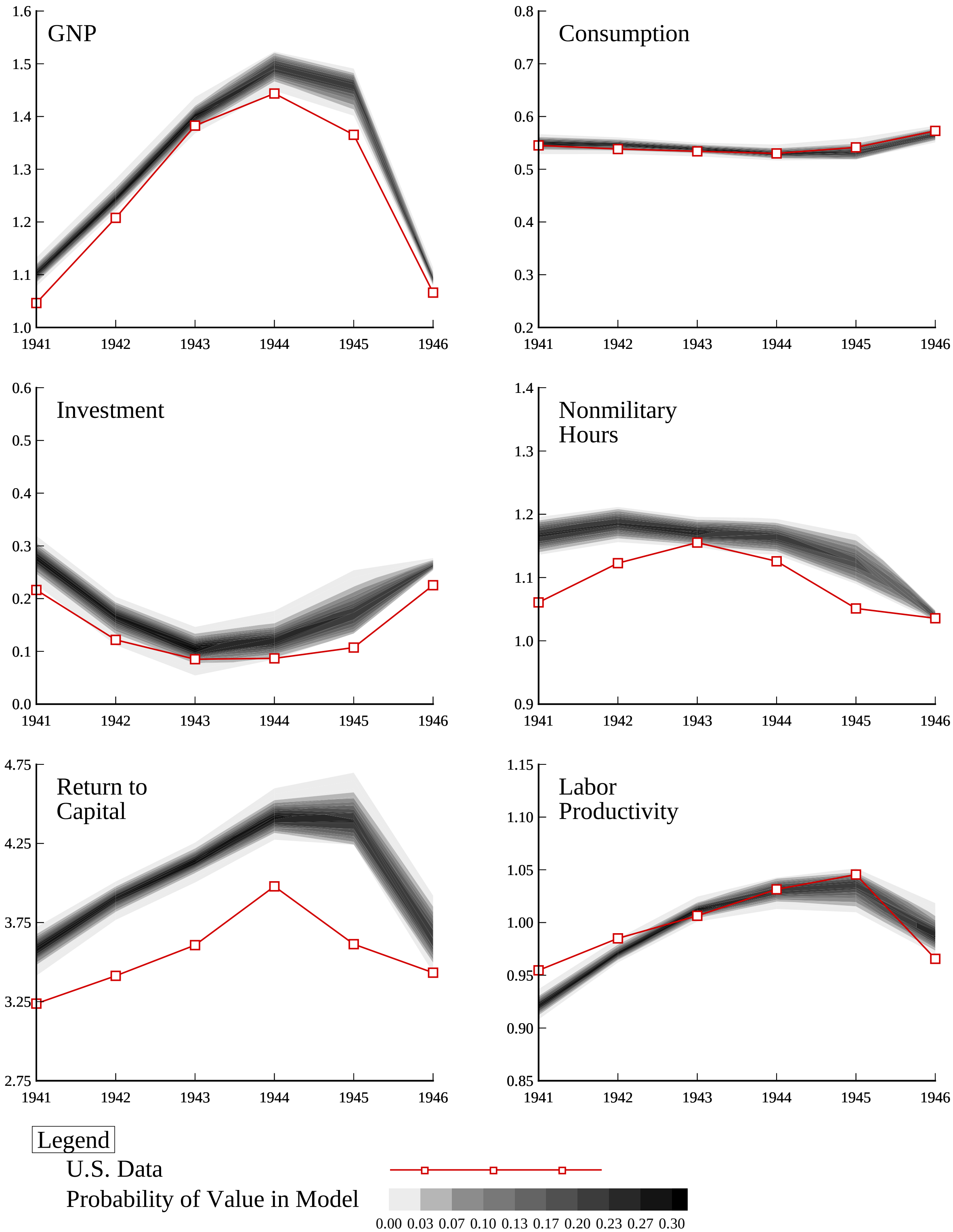
Figure 10. U.S. Data and Stochastic Model Predictions, 1941–1946
(Model with Labor Tax Rate Constant)



Legend
U.S. Data
Probability of Value in Model
0.00 0.03 0.07 0.10 0.13 0.17 0.20 0.23 0.27 0.30

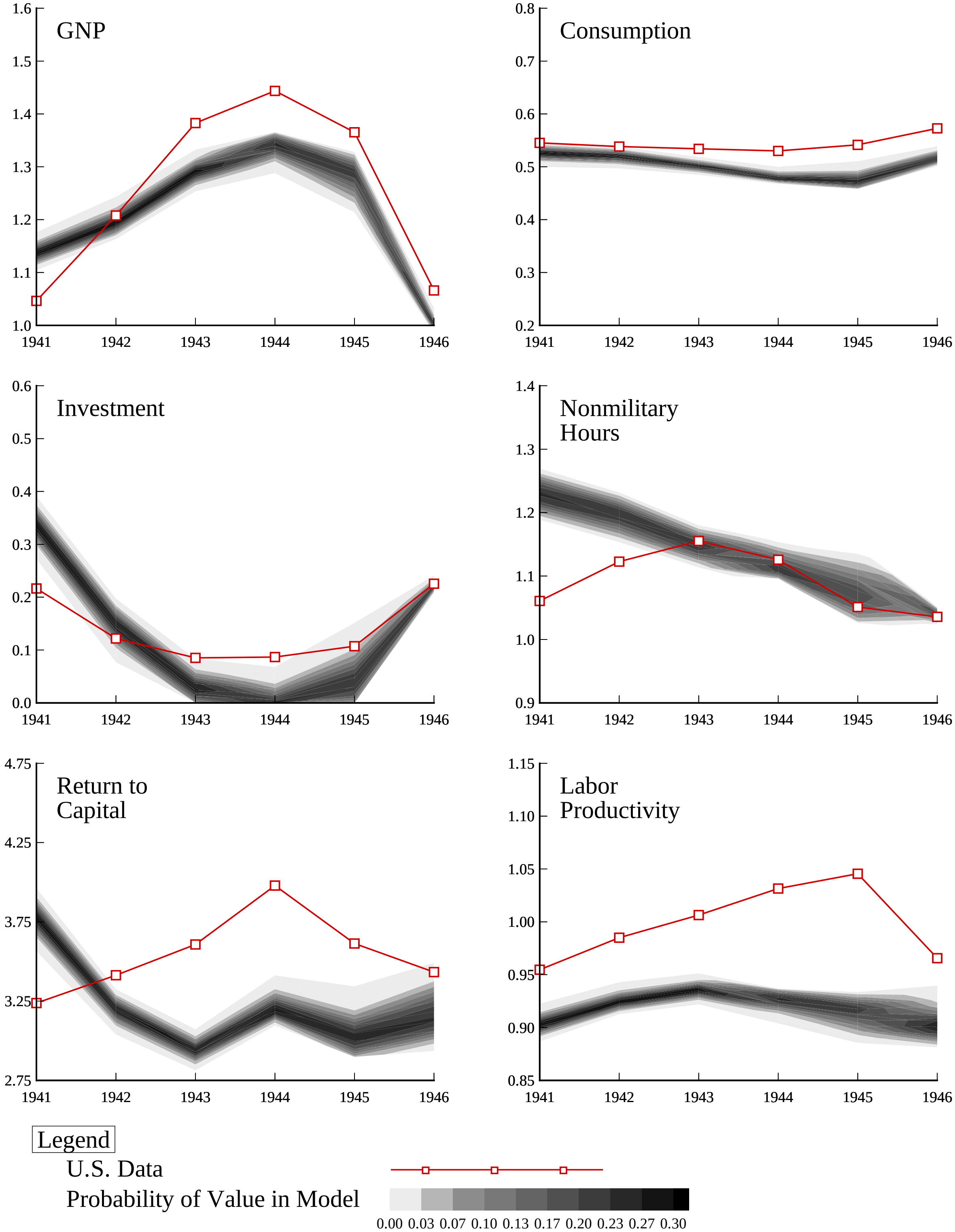
Note: See notes to Figures 5–7 for series normalizations.

Figure 11. U.S. Data and Stochastic Model Predictions, 1941–1946
(Model with Capital Tax Rate Constant)



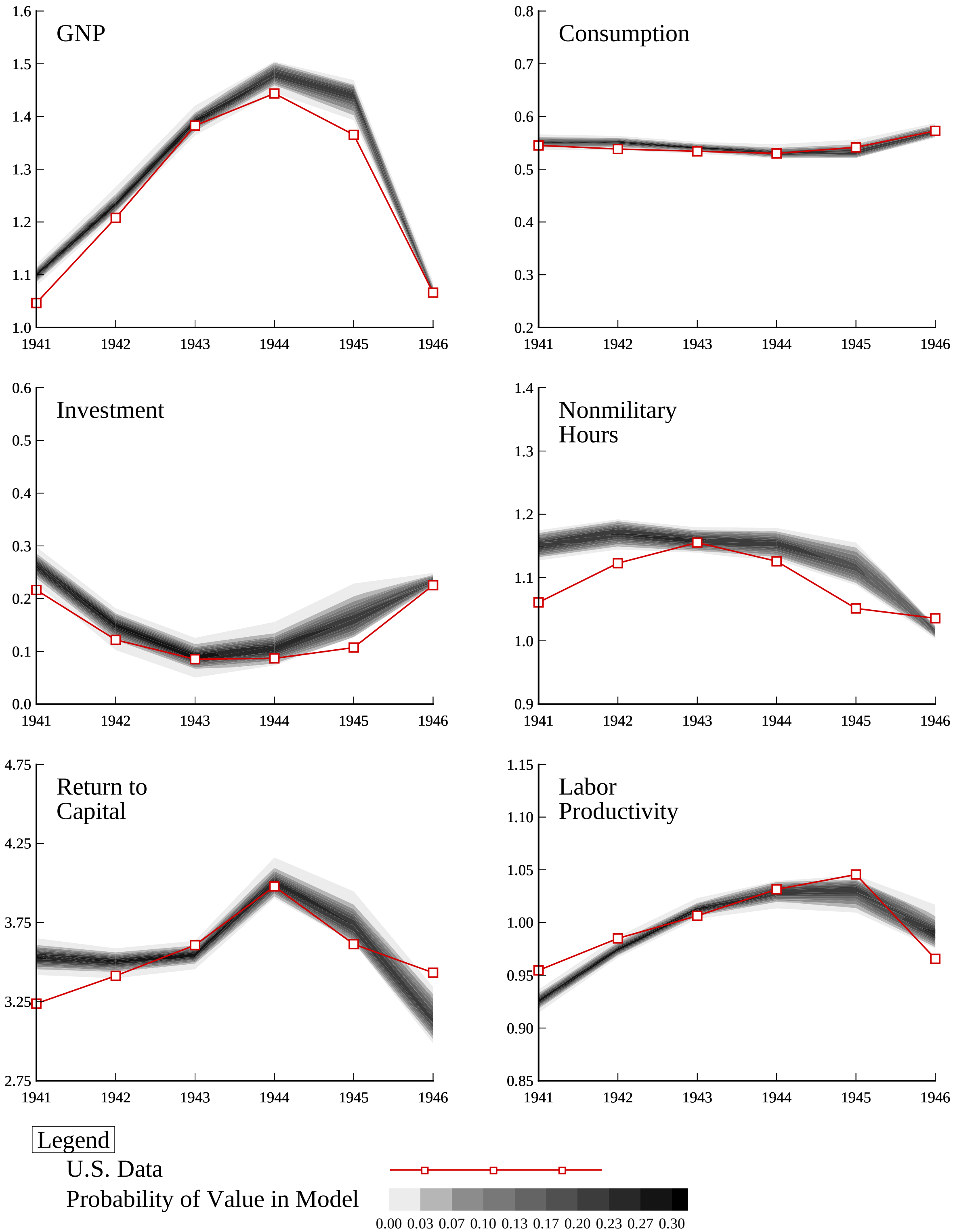
Note: See notes to Figures 5–7 for series normalizations.

Figure 12. U.S. Data and Stochastic Model Predictions, 1941–1946
(Model with Technology Constant)



Note: See notes to Figures 5–7 for series normalizations.

Figure 13. U.S. Data and Stochastic Model Predictions, 1941–1946
(Model with Postwar Depression State)



Note: See notes to Figures 5–7 for series normalizations.

Figure 14. U.S. Tax Rate on Labor and Tax Rate Needed in Model for No Change in Nonmilitary Hours of Work, 1941–1946