Chapter 22
Business Cycles

Mila Markevych and Alejandro Rojas-Bernal

Abstract This chapter surveys the evolution of business-cycle econometrics as a history of changing standards for empirical credibility. It traces recurring problems in the field: how to measure cycles, separate shocks from propagation, identify causal mechanisms, and use models for policy counterfactuals. Moving from early cycle theories and Cowles Commission structural models to rational expectations, RBC, VARs, DSGE, HANK, and production-network models, the chapter asks what has aged well and what has not. Its central lesson is that progress has come less from finding a final model of the cycle than from clarifying the assumptions needed for credible measurement, identification, and policy analysis.

22.1 Introduction

Business-cycle econometrics is best understood as a history of changing standards for empirical credibility, rather than as a sequence of models replacing one another. Each generation inherited a familiar set of problems: how to define the cycle, how to separate shocks from propagation, how to distinguish measurement from explanation, how to identify causal mechanisms from observational data, and how to use models for policy counterfactuals. What changed was the machinery used to answer these questions. Early cycle theorists worked with chronologies, turning points, and qualitative propagation mechanisms. The Cowles Commission converted business-cycle analysis into a problem of stochastic structure, simultaneity, and identification. Rational-expectations and equilibrium models redefined structure in terms of preferences, technologies, constraints, information, and policy rules. Vector

Mila Markevych ✉
University of North Carolina at Charlotte, Charlotte, NC, USA, e-mail: mila.markevych@charlotte.edu

Alejandro Rojas-Bernal
University of Hawai'i at Manoa, Honolulu, HI, USA, e-mail: rojasber@hawaii.edu

autoregressions (VARs) and related reduced-form methods recentered attention on the joint dynamics of the data. New Keynesian Dynamic Stochastic General Equilibrium (DSGE) models attempted to synthesize equilibrium discipline with nominal rigidities and Bayesian estimation. More recent Heterogeneous Agent New Keynesian (HANK) and production-network models have made heterogeneity, balance sheets, firm size, and supply chains part of the state of the economy itself. This chapter asks what, in that history, has aged well and what has not. The answer is not the same for every era. Some early substantive theories of cycles aged poorly: sunspot periodicities, mechanical long waves, deterministic regular cycles, and purely narrative accounts of crises are no longer adequate econometric explanations. Yet several of their organizing questions aged remarkably well.

The first theme that aged well is the separation between impulse and propagation as an organising modelling framework. The early literature already recognized that aggregate fluctuations are recurrent, economy-wide, and propagated through credit, investment, expectations, inventories, prices, and production linkages. Slutsky and Frisch gave this insight its durable econometric form by separating the impulse problem from the propagation problem. That distinction survived almost every later paradigm, becoming the first theme. It reappears in simultaneous equation models, VAR impulse responses, real business cycle (RBC) technology shocks, New Keynesian monetary transmission, financial-accelerator models, and heterogeneous-agent redistribution channels. The specific shocks changed, but the econometric problem did not.

A second theme concerns measurement. Business cycles are not directly observed objects. They are constructed from data through dating rules, detrending conventions, filters, stochastic decompositions, latent-state models, and real-time nowcasts. The National Bureau of Economic Research (NBER) tradition aged well in its insistence that cycles involve broad comovement across many series and that turning points require careful empirical documentation. Its weakness was the absence, in its early form, of a probabilistic model of how those patterns were generated. Later work improved on this by introducing stochastic trends, state-space methods, and dynamic factors. But the lesson was humbling: no measurement procedure is neutral. Hodrick-Prescott (HP) filters, band-pass filters, Beveridge–Nelson decompositions, unobserved-components models, Markov-switching regimes, and Hamilton-style forecast errors each define a different object. What aged well was the recognition that measurement must be explicit and reproducible. What aged poorly was the hope that one could extract *the cycle* mechanically before doing theory.

A third theme is identification. The Cowles program aged well in making identification a prerequisite for structural interpretation. Its core contribution was not any particular large-scale Keynesian model, but the idea that economic theory must impose restrictions on a probability law for observables, and that policy-relevant parameters cannot be recovered from data without assumptions. What aged less well was the confidence that large systems of behavioral equations, distributed lags, exclusion restrictions, and closure rules could deliver stable policy counterfactuals. The Lucas critique exposed one problem: estimated behavioral relations may shift when policy regimes change. Sims's VAR critique exposed another: many identifying restrictions in large structural models were neither transparent nor empirically compelling. Subsequent econometrics did not eliminate identification. It made the terms of identification more visible, through recursive and structural VARs, long-run and sign restrictions, narrative shocks, high-frequency surprises, external instruments, proxy SVARs, local projections, and state-space measurement.

A fourth theme is the persistent ambition to make business-cycle models useful for policy. This ambition aged well, however overconfidence in particular policy laboratories and methods did not. Large-scale macroeconometric models institutionalized forecasting and policy simulation, but often blurred empirical fit with structural credibility. RBC models restored equilibrium discipline and regime-invariant primitives, but asked too much of technology shocks, flexible prices, and representative-agent. New Keynesian DSGE models restored a role for monetary policy and nominal rigidities while retaining rational expectations and explicit equilibrium restrictions. Their success was real: they made structural models estimable, forecastable, and usable in central banks. Their limitation became equally visible after the financial crisis. A representative-agent, log-linear, frictionless model was not a sufficient laboratory for credit booms, balance-sheet crises, liquidity traps, deep unemployment, and unconventional policy.

The most recent turn beyond the representative agent continues rather than rejects this history. HANK and production-network models preserve the structural ambition of equilibrium macroeconometrics, but reject the shortcut that aggregate dynamics can usually be understood through a single household, a single firm, or a single shock. What has aged well is the demand for explicit mechanisms, disciplined counterfactuals, and coherent links between theory and data. What has aged poorly is the idea that greater formal detail by itself solves the old problems. Heterogeneity, networks, and computation add realism, but they also introduce new identification problems, new measurement choices, and new opportunities for misspecification. The history of business-cycle econometrics therefore offers a balanced lesson. Progress has come less from finding a final model of the cycle than from learning which claims require probability, which parameters require identification, which measurements require theory, and which counterfactuals require structural discipline.

22.2 Before Probability: Early Views of Economic Fluctuations (pre-1940s)

Before the probabilistic turn of the 1940s, economists had already developed concepts that remained central to business-cycle analysis. They shifted attention from isolated commercial crises to recurring economy-wide fluctuations; documented patterns of comovement, timing, and duration across sectors and time series; distinguished cyclical movements from secular trends; and framed cycles as the result of shocks transmitted and amplified through the economy. This section follows that development before the full probabilistic turn. It begins with early theories of recurring aggregate fluctuations, then turns to the descriptive measurement tradition, and finally traces the movement from regular-cycle theories to stochastic impulse-propagation frameworks, culminating in the first econometric cycle models.

22.2.1 Defining the Cycle: Early Theories

Business cycles were not always an object of theory in their own right. Through most of the nineteenth century, economists focused on crises as isolated commercial collapses, while the expansions that preceded them and the contractions that followed were treated as background to the main event. Juglar (1862) made the business cycle, rather than the crisis in isolation, the central object in the analysis of economic fluctuations. He recast commercial fluctuations as a recurrent, internally generated sequence of three stages: prosperity, crisis, and liquidation. First, *prosperity* is the phase of speculative expansion, in which credit growth finances rising investment commitments. Prices and deposits rise, while reserves decline. Second, the *crisis* is the turning point at which the investments made during prosperity are revealed to earn too little to repay the credit that financed them, leading confidence to break and credit to tighten. Finally, *liquidation* follows as the corrective phase, bringing bankruptcies, balance-sheet repair, debt repayment, and falling prices. Liquidation restores the conditions for renewed expansion and a new phase of prosperity.

By contrast, Jevons (1879) accepted the recurrence of commercial crises but sought to explain it through an exogenous periodic force. In his account, solar cycles affect climate and harvests, and he treated sunspots as a proxy for solar activity. Harvest fluctuations then propagate into prices, incomes, trade, credit, and speculation, so that crises appear as the collapse phase of a broader trade cycle.

Wicksell (1898) transformed the analysis of cycles into a monetary theory centered on the gap between the *market interest rate* set in credit markets and the *natural interest rate* consistent with the economy's underlying saving-investment balance. When the market rate falls below the natural rate, firms find it profitable to borrow because the expected return on investment exceeds the cost of finance. Credit, deposits, and spending then expand. Because productive capacity, prices, and wages do not respond immediately, this expansion puts upward pressure on prices and profits, sustaining both credit demand and the cumulative process. This disequilibrium ends only when the interest-rate gap is closed. Banks may raise the market rate in response to inflation, reserve pressure, convertibility constraints, balance-sheet fragility, or rising default risk. The natural rate itself may move in response to changes in saving and investment fundamentals. This Wicksellian mechanism became a central point of departure for the Austrian theories of Mises (1912) and Hayek (1929, 1931). Mises turned it into a theory of bank-created credit and unsustainable investment plans, while Hayek recast it as a problem of intertemporal misallocation in which artificially cheap credit distorts the structure of production across the economy. More broadly, these authors redefined the cycle as a deviation from an underlying intertemporal equilibrium, yet they differed sharply in policy implications. For Wicksell, the interest-rate gap could in principle be closed by corrective monetary policy, whereas for Mises and Hayek, the recession was the necessary correction of distortions created during the boom.

Schumpeter (1939) characterizes the business cycle as an endogenous feature of capitalist development driven by innovations implemented by entrepreneurs and financed by bank credit. In his account, a primary wave begins when new innovations redirect resources, raise investment, and generate temporary profits. A secondary wave then follows as established firms react, credit expands more broadly, and speculative behavior amplifies the boom. Because these induced responses can dominate the aggregate data, the initiating role of innovation may be hard to see in observed fluctuations. Recession is therefore not a purely accidental breakdown, but an adaptation phase in which the economy reorganizes itself around the new production structure. Obsolete firms are eliminated or forced to adjust, and profits and balance sheets come under pressure.

Schumpeter later classified cyclical movements into Kitchin, Juglar, and Kondratiev cycles, treating this classification as a useful descriptive device. Kitchin (1923) cycles are short inventory-driven fluctuations of roughly 3–5 years. Juglar cycles are medium-run fluctuations of roughly 7–11 years, usually associated with fixed investment. Kondratiev (1935) cycles are roughly 48–60 years and are often linked in Schumpeterian narratives to broad technological eras.

Fisher (1933) proposed a theory of crisis amplification and depression dynamics in which over-indebtedness, once combined with deflation, sets off a cumulative debt-deflation spiral. Debt liquidation forces distress selling and loan repayment, which shrinks bank deposits and slows the pace at which money changes hands. Prices then fall, raising the real burden of outstanding debts. As real debt commitments rise, net worth and profits collapse, bankruptcies spread, investment and output contract, and pessimism intensifies, further increasing hoarding and reinforcing the downturn. In Fisher's account, the crisis becomes self-aggravating because the attempt to reduce debt through liquidation depresses prices so sharply that the more debtors pay, the more they owe in real terms. This diagnosis led Fisher to advocate reflation of the price level as the central policy response, since only reversing deflation could break the spiral of rising real debt burdens and a collapsing demand.

Keynes (1936) built a theory of the cycle centered on investment-driven fluctuations in effective demand, the total spending that determines output and employment. Investment depends on the marginal return of capital, which is highly sensitive to unstable long-term expectations, confidence, and what Keynes called animal spirits. Because employment and income are determined by effective demand, underemployment can persist whenever investment is too weak to sustain full-employment output. For Keynes, wage rigidity could worsen a downturn, but it was not the fundamental problem. The larger issue was that, under uncertainty, interest rates could remain too high relative to expected investment returns, while wage cuts might do little to restore employment and could further damage expectations and confidence. Keynes therefore defended monetary easing and interest-rate management as useful, but often insufficient, remedies. In severe downturns, fiscal expansion becomes necessary to stabilize demand and secure full employment.

As the discussion turned from theory to measurement, a common problem came into view: how to distinguish short-run fluctuations, often driven by credit conditions, expectations, or effective demand, from slower-moving structural forces. For Wicksell, cyclical movements came from deviations between the market and natural rates. For Schumpeter, they took the form of innovation-driven development and their secondary-wave effects. For Keynes, they were shortfalls of effective demand relative to full-employment output. The task of measurement required more than simply isolating 'the cycle'. It also required specifying the underlying benchmark – trend, equilibrium, or evolving structure – against which cyclical deviations are to be defined.

22.2.2 The National Bureau of Economic Research (NBER) Program and Business-Cycle Measurement

As a leading figure in American institutional economics, Mitchell (1913, 1927) treated business cycles as a historically feature of the modern money- and profit-driven economy, composed of many interconnected yet imperfectly synchronized processes – production, prices, investment, inventories, profits, and employment. This complexity implied that no single series, equation, or mechanism could adequately capture the cycle. At the NBER, which Mitchell shaped as its founding Director of Research, the task became to document business-cycle chronologies, turning points, phases, and co-movements across many time series rather than to reduce fluctuations to a single causal law.

This era of measurement was characterized by the continual appearance of partly mechanical and partly judgmental methods for smoothing, detrending, and isolating cyclical movement. Economic time series were viewed through an approximate decomposition of the form $Y_t \approx T_t + S_t + C_t + I_t$, where Y_t is the observed series, T_t a long-run trend, S_t seasonal variation, C_t cyclical movement, and I_t irregular disturbances. In this framework, trend was usually conceived not as a stochastic process but as a smooth deterministic path, sometimes fitted by a line or curve, sometimes even drawn freehand through years taken to represent relatively normal or 'active but unhurried' production.

This era's mentality is evident in debates over whether to remove seasonality before or after detrending, since the order could change the measured cycle. Kuznets (1930) extended this measurement tradition by making secular trend itself an object of quantitative analysis. Rather than treating trend merely as something to remove before studying cycles, he measured long-run movements directly and examined how they interacted with medium-run swings and shorter cyclical fluctuations in production and prices. Burns (1934) pushed this further with overlapping 'working decades' to separate secular trend from shorter-run cycles, and broadened the cycle to include irregular disturbances.

The experimental character of this measurement era is visible even in the evolution of the NBER chronology itself. As Romer (1994) later showed, before 1927, definitions of the NBER turning points relied heavily on detrended business indexes, whereas later definitions were based much more on levels of economic activity. The earlier approach tended to date peaks too early and troughs too late, thereby making prewar recessions appear longer than they were under later dating conventions. In the postwar NBER tradition, Burns and Mitchell (1946) systematized business-cycle measurement by formalizing concepts such as business-cycle chronology, reference and specific cycles, turning points, durations, amplitudes, conformity, and the role of seasonal adjustment in comparing cyclical patterns across many series. Moore (1961) complemented these measurement strategies with real-time macro-monitoring exercises aimed at identifying where the economy stands in the cycle, which processes tend to lead or lag, how cyclical movements diffuse across sectors, and how current data can be used for diagnosis and short-term forecasting.

The measurement tradition excelled at documenting cycles, and its descriptive tools have remained central to empirical business-cycle analysis. But it lacked an explicit probabilistic and structural account of how observed fluctuations were generated. Yule (1927) already pointed to the problem by arguing that an irregular cyclical series need not be a fixed periodic law masked by random noise. Instead, an irregular cycle could be generated when random shocks pass through a system with recurring internal dynamics. The resulting movements may look smooth and cyclical, but their amplitude and timing shift over time as new disturbances enter the system.

At this point, measurement alone was no longer enough. Economists had to move beyond the search for hidden regular cycles toward models of impulse and propagation: models that explained how shocks entered the system and how internal dynamics transformed them into observed fluctuations.

22.2.3 From Regular Cycles to Impulse-Propagation

Slutsky (1927, 1937) challenged the inference from regular cyclical patterns to genuinely periodic causes, showing that random disturbances, once transmitted through a systematic mechanism, can generate wave-like dynamics. He showed this using a moving-average representation. If each observation is formed from a weighted sum of current and past random shocks, then adjacent observations share many of the same components. This mechanical overlap creates persistence and smoothness in the resulting series. The series may therefore appear cyclical even though no periodic force has been introduced into the underlying shocks.

Frisch (1933) introduced this critique into a macroeconomic framework by distinguishing the *impulse problem* – how shocks enter the system, from the *propagation problem* – how they move through its internal structure. A simple accelerator (investment responding to consumption and its rate of change) converges smoothly rather than oscillates, so Frisch introduced propagation through his carry-on-activity model in which production already underway shapes current activity, creating a delay structure that transmits, dampens, and transforms shocks into cyclical movements. This marked a decisive move from descriptive chronology to probabilistic dynamics.

22.2.4 The First Econometric Cycle Models

Tinbergen (1939a, 1939b, 1940) turned the impulse-propagation program into an estimable macroeconomic system. His methodological break was to specify a set of structural equations whose variables and lags were chosen in advance on the basis of economic reasoning. The aim was not to reproduce every detail of the economy, but to build a complete system in which each endogenous variable was explained by a behavioral, definitional, institutional, or accounting relation. Cyclical movements therefore emerged from the interaction of linked equations, rather than from descriptive co-movement among series alone. In Tinbergen's own language, business-cycle verification had to proceed through "direct causal relationships," and econometric research had to combine mathematical economic structure with statistical measurement.

Tinbergen's simple examples made the point clear. Investment was modeled as responding to past profits, consumption as responding to wages and past profit movements, and profits as determined by the accounting relationship between consumption, investment, and wages. Once these equations were linked together, the estimated coefficients did more than summarize empirical correlations: they determined whether the system converged smoothly, oscillated, or became unstable over time. Tinbergen applied the same logic in a larger model built for the League of Nations, reducing the system to 'final equations' that summarized the model's implied cyclical period and damping.

Yet the method was still incomplete. Tinbergen estimated many behavioral relations equation by equation, even though the model itself implied simultaneity. He did not yet possess a full probability theory for identification and inference for simultaneous systems. Keynes's criticisms targeted precisely this weak point. He argued that Tinbergen's results could be distorted by omitted forces such as expectations and confidence, multicollinearity, lag selection, trend handling, linear approximation, and coefficient instability across periods. More generally, Keynes worried that statistics might ratify theory *ex post facto* rather than discipline it. Hence his famous remark that Tinbergen was the person he would trust with "black magic", while denying that the method was yet a science. The importance of the controversy is that Tinbergen's revolutionary structural program exposed exactly what was yet to be added before macroeconometrics became methodologically mature.

Seen from the present, the enduring value of this era was conceptual rather than fully explanatory. It supplied the facts to be explained, the distinction between trend and cycle, and the impulse-propagation framework that later macroeconomics retained. But it could not yet provide a probabilistic account of shocks, inference, or structural counterfactuals.

What did not age well was the absence of stochastic foundations. Deterministic cycle theories could narrate recurrent motion, but without stochastic foundations they could not distinguish systematic mechanisms from chance regularities. Description without probability allowed chronology and pattern recognition, but it provided weak grounds for inference, no credible counterfactuals, and no disciplined way to assess uncertainty. Likewise, detrending practices could separate apparent long-run movement from short-run fluctuation, but when carried out mechanically they offered no model of persistence and no account of why residual movements should be interpreted as ‘cycles’ rather than noise or misspecification. Most fundamentally, the early literature lacked identification: shocks were not clearly separated from the mechanisms that propagated them through the economy. These gaps are precisely what made the next revolution necessary: probabilistic structural models, explicit identification, and estimation methods designed to separate disturbances from propagation and to turn business-cycle analysis into a more disciplined empirical science.

22.3 The Structural Ambition: Macroeconometrics After the Cowles Commission (1940s-1970s)

The move from early business-cycle theory to postwar macroeconometrics required more than adding regressions to economic models. It required a new way of thinking about what it meant to estimate an economic structure from observational data. That methodological foundation was supplied largely by the Cowles Commission, the research group that formalized the structural-equation approach to econometrics. Its central contribution was to distinguish structural equations from reduced-form relationships, to make identification a prerequisite for estimation, and to treat economic theory as a probabilistic system whose parameters could be tested, estimated, and used for counterfactual policy analysis. The decades that followed marked the high point of macroeconometrics’ early *structural ambition*. The goal was no longer merely to describe business-cycle regularities or to fit aggregate time series, but to recover policy-relevant economic structure from data. This section follows that ambition from the simultaneous-equation framework of the 1940s and 1950s, through the intermediate quarterly models of the early 1960s, to the large-scale macroeconometric systems that came to dominate postwar forecasting and policy work.

22.3.1 Simultaneous Equation Models

Haavelmo (1944) shaped modern econometrics by arguing that an economic theory becomes econometrically meaningful only when translated into restrictions on the probability law of observable data. An exact theoretical relation $f(x^* \mid \theta) = 0$ can be tested on raw observations only if theoretical variables coincide exactly with measured magnitudes and there is no error, implausible in social data, so the theory must be reformulated in stochastic form $f(x \mid \theta) = u$, with the disturbance treated as part of the model rather than as an extraneous residual. Econometrics then specifies the implied probability law for observables along with the distributional assumptions governing u, and judges a model by whether the observed sample can reasonably be treated as a realization from that law rather than by point-by-point fit.

The structural econometric model then attempts to recover from passive observation an autonomous economic relation consistent with theory. In doing so, it faces three classic problems. *Reversibility* asks whether the available data are the right kind of evidence to recover the theoretical relation. *Simplicity* asks whether the model remains parsimonious without losing its ability to fit the facts. Finally, *autonomy* asks whether the relation would remain valid under changes in policy, institutions, expectations, or other structural conditions. These three difficulties become especially acute once the economy is represented not by a single stochastic relation but by a simultaneous system of structural equations. In this case, the data reflect equilibrium outcomes jointly determined by several interacting structural relations and cannot be treated as evidence of a single autonomous behavioral relation.

To make this clear, consider the *structural system* $Ay_t + \Gamma z_t = u_t$, where y_t is the M vector of endogenous variables, z_t – the vector of predetermined variables, including exogenous variables and lagged endogenous variables, and u_t – the vector of structural disturbances, with distribution summarized, at a minimum, by its covariance matrix Σ. Haavelmo's contribution was to insist that the structural equations and the stochastic specification must be treated jointly, and the model should map parameters such as (A, Γ, Σ) into a joint probability law $P(y, z)$ for the observables. He insisted that it is generally invalid to estimate the equations one by one as if each were statistically separable from the rest of the system, because they hold simultaneously and therefore impose joint restrictions on the observable data.

Koopmans (1949, 1950) generalized Haavelmo by confronting the *identification problem*: under what restrictions can the structural parameters be recovered uniquely from the reduced-form relationships? In a simultaneous-equation model, predetermined variables *excluded* from a given equation are the source of identification because they can shift the endogenous right-hand-side variables without entering the equation directly. The resulting *order condition* (the number of excluded predetermined variables must be at least as large as the number of endogenous regressors) is necessary. The *rank condition* (those excluded variables must generate enough independent reduced-form variation in the endogenous regressors – formally, full row rank of the relevant reduced-form block) is necessary and sufficient in the standard linear case.

After identification came the problem of estimation. In a simultaneous system, equation-by-equation ordinary least squares is generally inconsistent because the endogenous right-hand-side variables are jointly determined with the dependent variable and are therefore correlated with the structural disturbance. The Cowles contribution was to replace naive single-equation regression with estimators derived from the stochastic structure of the whole system. In the linear-Gaussian case this led to full-information maximum likelihood (FIML). T. W. Anderson and Rubin (1949) developed limited-information maximum likelihood (LIML) for individual structural equations. Estimation became inseparable from identification: one first had to show that a structural equation was recoverable from the reduced form, and only then ask how to estimate its coefficients in a statistically coherent way.

By the early 1950s, simultaneous-equation estimation moved toward computationally simpler projection-based methods. Theil (1953) introduced two-stage least squares (2SLS), and Zellner and Theil (1962) extended the logic to system-wide three-stage least squares (3SLS), which uses the estimated cross-equation covariance matrix to gain efficiency when structural disturbances are correlated across equations.

Klein (1947, 1950); Klein and Goldberger (1955) turned the Cowles framework into operational Keynesian macroeconometrics, progressing from a provisional postwar prototype to a systematic structural program to the institutionalized Klein–Goldberger national model. Taken together, these works established the template for postwar structural macroeconometrics that bridged early Cowles–Keynesian models and the large national systems of the 1960s.

22.3.2 Large-Scale Macroeconometric Models

The transition from the early annual Cowles–Keynesian systems to the large quarterly models of the mid-1960s passed through an intermediate stage that combined quarterly data with the simulation of dynamic adjustment paths rather than static multipliers. Duesenberry, Eckstein and Fromm (1960), T.-C. Liu (1963), and Klein (1964) brought recursive recession simulation, nonrecursive effective-demand systems, and forecasting-oriented inventory and capacity blocks to the field, exposing both the value of dynamic quarterly simulation and the fragility of structural estimation in small interdependent systems.

The next step was a change in scale. Large-scale modeling required common datasets, consistent accounting definitions, block structure, closure rules, and numerical solution methods able to produce forecasts, multipliers, and counterfactual paths for the economy as a whole, transforming macroeconometrics into an institutional technology for policy analysis.

Large-scale macroeconometric modeling therefore did not evolve along a single path, but through several strategies for making quarterly structural models usable. Dusenberry, Fromm, Klein and Kuh (1965) represents the system-engineering route. The Brookings–Social Science Research Council (SSRC) model was a collaboratively built system of more than 150 structural equations and 75 identities, organized largely in block-recursive form, with earlier blocks feeding later ones but with a simultaneous core for consumption, inventories, orders, labor requirements, and factor incomes. Its contribution was to shift attention from isolated behavioral equations to the coordination, refitting, and iterative solution of the whole model, so that the system could generate consistent forecasts, economy-wide multipliers, and policy simulations. De Leeuw and Gramlich (1968) represents the monetary-transmission route. The Federal Reserve–MIT model retained the large block structure but placed financial markets at the center, linking reserves, deposit demand, bank credit, portfolio adjustment, the term structure, and mortgage rates to housing, fixed investment, inventories, consumption, and spending. Because its price–wage–labor block was still incomplete, its main contribution was less full macroeconomic closure than a more explicit Keynesian account of monetary propagation. Duggal, Klein and McCarthy (1974) represents the mature operational forecasting route. Wharton Mark

III recast the large quarterly model as an integrated IS–LM system, linking product, factor, and financial markets within a single forecasting apparatus and evaluating the model through dynamic simulations, forecast-error analysis, cross-correlation diagnostics, and fiscal and monetary counterfactuals.

The three traditions shared equation-by-equation estimation, distributed lags, and simulation-based policy analysis, but emphasized different central problems: Brookings – solution of a huge nonlinear system; Fed–MIT – monetary-transmission channels; Wharton – operational forecasting. These tradissions then diffused institutionally through the OBE/BEA, commercial vendors (WEFA, DRI), and the international Project LINK, transforming Cowles–Keynesian macroeconometrics into a practical infrastructure for forecasting and policy simulation.

22.3.3 Policy Counterfactuals and Emerging Fragilities

Once large-scale macroeconometric systems were in place, the central promise of macroeconometrics shifted. The main question shifted from whether one could fit historical relationships or generate a baseline forecast to whether one could use an estimated structural system to conduct counterfactual policy analysis. Policymakers wanted more than projections of output and prices under unchanged conditions. They wanted explicit answers to conditional questions: What would happen if taxes were cut? If defense spending rose? If reserves were tightened? If mortgage rates increased? If exports weakened? Or if income policies were imposed? The key objects of analysis became impact multipliers (i.e., the immediate effect of a policy change), dynamic multipliers (i.e., policy effects accumulated over time), and the full time paths of the economy under alternative policy scenarios. Large-scale models thus turned macroeconomics into a tool for studying how output, employment, inflation, and the balance of payments would evolve under different assumptions about policy and the external environment.

In practice, policy counterfactuals meant solving the model twice, once along a baseline, then again after changing a policy instrument, and reading the difference as the estimated effect. Macroeconometrics thereby evolved into an instrument for constructing conditional policy histories.

This counterfactual methodology was convincing because it appeared to combine empirical discipline with institutional realism. Large-scale models encoded detailed knowledge of fiscal programs, monetary instruments, financial markets, and sectoral spending. They decomposed aggregate demand into policy-relevant components and generated quantitative scenarios rather than qualitative predictions. Compared with single-equation reduced forms, they seemed richer because they described multiple policy channels: how taxes affected disposable income and consumption; how reserves affected interest rates, deposits, housing, and business investment; and how government purchases affected output, inventories, and prices. This made them attractive to policy agencies. Beyond forecasting, they served as operational tools for budget planning, monetary analysis, and recurrent policy simulations.

Yet their success as counterfactual devices also made their weaknesses harder to hide. As large-scale models became tools for comparing policy scenarios, their results depended increasingly on assumptions about expectations, wage and price adjustment, and financial-market behavior. A model could forecast tolerably well in normal conditions and still be fragile for policy evaluation if behavioral coefficients shifted with institutions, if key sectors were thinly modeled, or if important channels, especially prices, finance, and international spillovers, were only partially captured. This raised a deeper question: could systems useful for forecasting also deliver structurally credible counterfactuals? Four fragilities became especially important.

Their success as counterfactual devices made their weaknesses harder to hide. First, the Cowles problem of autonomy reappeared at the scale of policy simulation, since coefficients estimated under one policy and institutional environment could forecast tolerably yet fail to remain stable when rules, expectations, or institutional constraints changed. Second, the systems increasingly depended on pragmatic identifying restrictions, distributed lags, and closure rules that weakened the structural interpretation of their counterfactuals, so a solved simulation could look precise while resting on only partly justified assumptions.

The most visible substantive weaknesses concerned prices, inflation, and finance. Price-and-wage equations were among the least secure parts of the systems, and financial blocks required repeated respecification – the Federal Reserve–MIT/Penn model already ran policy simulations without a fully endogenous account of inflation and wage formation, and Wharton Mark III identified price forecasting, import prices, and capacity constraints as areas needing revision. These weaknesses sharpened the tension between forecast performance and structural credibility, and Brunner (1973) argued that pragmatic data-tracking through intercept adjustments and dummy variables did not validate the theory or policy conclusions built into the model.

Against this background, the St. Louis economists L. C. Andersen and Jordan (1968); L. C. Andersen and Carlson (1970) recast macroeconometric policy analysis as a smaller reduced-form test of policy effectiveness, with monetary and fiscal actions entering as summary policy variables. Their evidence pointed to a stronger role for monetary actions and made the point that smaller reduced-form models stated more clearly what was being tested and were less exposed to the fine-tuning that had weakened the empirical content of large-scale systems.

By the early 1970s, the Cowles–Keynesian structural program had produced a genuine methodological advance and a visible impasse. Its lasting contribution was to give macroeconometrics a language of probability, identification, estimation, and counterfactual policy analysis. It showed that theory becomes empirically meaningful only when it imposes restrictions on observables, and it made policy simulation a central empirical task. Yet the same program became fragile as its models grew larger and more policy-relevant. Identification often rested on exclusion restrictions that were only partly convincing; dynamics were frequently introduced through ad hoc lags and expectation terms; and price formation, financial propagation, and policy-regime change were modeled with less theoretical discipline than the rest of the system. The result was a split in the subsequent literature. One response sought deeper policy-invariant structure in rational-expectations and equilibrium business-cycle models. The other moved toward more transparent reduced-form time-series methods. Both responses preserved part of the Cowles ambition, but neither accepted the large-scale Keynesian model as a stable representation of economic structure.

22.4 The Equilibrium Turn in Business-Cycle Theory (1970s–1980s)

The equilibrium turn of the 1970s and 1980s was the most important response to the perceived weaknesses of postwar Keynesian macroeconometrics. Its central claim was that aggregate relations could not be treated as stable behavioral equations independent of the expectations, information sets, and policy rules under which they were formed. Expectations were part of the equilibrium itself, rather than extra variables to be tacked on through mechanical lag rules. Policy instruments were parts of rules that private agents observed, forecasted, and built into their decisions, instead of arbitrary paths to be fed into a fixed system. This changed the meaning of structure in macroeconometrics. The meaning of a structural model expanded from a simultaneous system identified by exclusion restrictions to a full specification of preferences, technologies, constraints, information, policy rules, and the equilibrium decision rules they generate.

This era gave rise to a new discipline for macroeconometric reasoning. Business-cycle theory had to be written as a stochastic equilibrium system. Shocks had to be identified, frictions had to be explicit, measurement had to be theory-aware, and model evaluation had to confront uncertainty rather than rely only on selected moment comparisons. This section follows that development from rational expectations and the Lucas critique, through real business-cycle theory and calibration, to the empirical tensions that necessitated subsequent methodological advances. These would eventually take the form of Dynamic Stochastic General Equilibrium (DSGE) modeling.

22.4.1 Rational Expectations and the Critique of Policy Analysis

The rational-expectations revolution did not begin as a business-cycle theory, but as a restriction on how agents formed forecasts. Agent's expectations had to be determined inside the same model that described prices, quantities, and information instead of being added afterwards through adaptive lags, extrapolations, or survey proxies. The policy implication was direct: when a policy rule changes, agents revise their forecasts and decision rules, so historically estimated relationships may no longer be reliable for counterfactual analysis.

Muth (1961) defined rational expectations as model-consistent forecasting: agents' beliefs about future variables match, on average, the probability distribution implied by the model. Expectations therefore became forecasts generated jointly with the model's own laws of motion, rather than adaptive distributed lags appended to behavioral equations. Forecast errors could occur but not be systematically predictable from information already available to agents.

Rational expectations entered business-cycle theory through the natural-rate and Phillips-curve debates of the late 1960s and early 1970s. Phelps (1967) and Friedman (1968) rejected the idea that the Phillips curve offered policymakers a stable long-run tradeoff between inflation and output. In reduced form, the natural-rate hypothesis says that output deviates from its natural level only when actual inflation differs from expected inflation, apart from real disturbances. If policy raises inflation unexpectedly, output may rise temporarily. Once expected inflation adjusts, the real effect disappears and the economy returns to its natural level.

Lucas (1972) gave the natural-rate logic a rational-expectations and market-clearing foundation: agents observe local prices but not the aggregate state, so they sometimes mistake nominal disturbances for relative-price signals and temporarily change production. Monetary shocks therefore have short-run real effects, and an econometrician may estimate a positive output–inflation relation, but conditional on the information structure rather than as a stable policy tradeoff. Lucas (1973) then drew the cross-regime implication: the slope of the observed Phillips curve should be steeper where nominal demand is volatile (agents attribute more price movement to aggregate shocks) and flatter where it is stable, so the estimated output–inflation tradeoff is not a structural parameter fixed across regimes.

Sargent and Wallace (1975) extended the rational-expectations critique to the choice of monetary-policy instrument. In their model real activity responds to inflation surprises, and the expected inflation is determined by the announced policy rule rather than by an adaptive lag. The predictable component of a money-supply rule is therefore absorbed into private forecasts and cannot create the inflation surprise needed to move real output, while only the unpredictable innovation can. A nominal-interest-rate peg, by contrast, fixes the rate while accommodating any demanded quantity of money, leaving the nominal scale and price level indeterminate. The broader lesson is that apparent policy tradeoffs are sensitive to the expectations assumption: they arise under adaptive expectations but disappear under rational expectations.

Barro (1976) connected the Lucas imperfect-information story with the Sargent–Wallace policy-ineffectiveness result by redefining the policy benchmark: policy should move the economy toward the full-information allocation rather than hold output at a fixed target. This leaves only a narrow role for activist monetary policy, since monetary volatility is harmful precisely because it makes prices less informative, and feedback rules based only on public information are neutral. Countercyclical policy can help only when the monetary authority has current information private agents lack, and even then, direct disclosure does the same work as intervention.

Lucas (1976) stated the central methodological critique: large-scale macroeconometric models could forecast tolerably under an unchanged regime but fail at counterfactual policy evaluation if they treated historically estimated decision rules as invariant when the regime changed. Writing the economy as $Y_{t+1} = F(Y_t, x_t, \theta, e_t)$, the conventional approach holds the behavioral parameters θ fixed while varying the policy variable x_t. Lucas's point was that θ is not a deep primitive – it summarizes expectations and decision rules formed under a particular policy environment. When the policy is itself a rule $x_t = G(Y_t, \lambda, \eta_t)$ with regime λ, the behavioral coefficients depend on the regime: $\theta = \theta(\lambda)$. Policy evaluation must therefore compare explicit policy rules and derive the equilibrium induced by each rule, not change x_t with θ fixed.

Fischer (1977) and Phelps and Taylor (1977) showed the policy-ineffectiveness result is fragile under nominal commitments: with overlapping wage contracts or preset prices a known feedback rule can affect short-run output even with rational expectations, so policy effects must be derived from explicit frictions, information sets, and rules rather than from historically estimated Phillips curves.

Kydland and Prescott (1977) added a time-consistency point: a plan that looks optimal before private expectations are formed may no longer be optimal once private decisions are fixed, so agents may not believe the original plan. In the inflation–unemployment example the temptation to create surprise inflation is anticipated, yielding an inflation bias without lower unemployment. Policy analysis therefore shifts from optimal discretionary sequences to credible rules in equilibrium.

By the late 1970s, Lucas and Sargent (1979) turned the rational-expectations critique into a broader case for equilibrium business-cycle analysis: stochastic models with optimizing agents, rational expectations, and complete laws of motion rather than semi-autonomous behavioral equations, providing the immediate background for real business-cycle theory. L. P. Hansen and Sargent (1980) made the rational-expectations program operational for dynamic linear models: rather than estimate private decision rules as distributed lags, the econometrician estimates the model's primitives (preferences, constraints, shock processes) and derives decision rules from optimization and rational expectations. In linear–quadratic settings this ties expectations, decision rules, and observable dynamics together through cross-equation restrictions on the coefficients of the time-series representation, providing the econometric bridge from the Lucas critique to equilibrium business-cycle and DSGE models.

22.4.2 Real Business Cycle Models and the Rise of Calibration

Real business-cycle (RBC) theory turned the Lucas critique into a positive research program. If policy evaluation required parameters stable across regimes, macroeconomists had to build models from preferences, technologies, constraints, information, and equilibrium conditions, and derive the economy's stochastic behavior from those primitives. The object of analysis shifted from reduced-form correlations to artificial economies: dynamic environments in which optimizing agents respond to shocks through equilibrium decision rules. Business cycles were therefore studied as equilibrium fluctuations of a neoclassical economy hit by real disturbances.

This shift also changed empirical practice. Instead of estimating separate consumption, investment, labor supply, wage, price, and money demand equations, RBC researchers specified a complete stochastic general equilibrium model, solved and simulated it, and compared its moments with those of aggregate data. Calibration supplied the practical method: parameters were chosen from steady-state restrictions, long-run ratios, microeconomic evidence, or independently estimated elasticities, and the model was judged by its ability to reproduce the volatility, persistence, and comovement of major aggregates such as output, consumption, investment, hours, and productivity. This was quantitative macroeconomics, but not macroeconometrics in the Cowles–Keynesian sense. The central question was whether a disciplined artificial economy could generate the main business-cycle facts, not whether each behavioral equation fit the data separately.

Kydland and Prescott (1982) supplied the first canonical implementation of the real business-cycle program by embedding aggregate fluctuations in a neoclassical growth model with two mechanisms: stochastic technology shocks and time to build. Technology shocks changed expected returns, labor supply, consumption, and investment; time to build spread the effects of current investment over future productive capacity. The model therefore reinterpreted Frisch's impulse–propagation tradition inside a competitive equilibrium economy: the impulse was a real productivity shock, proxied by the Solow residual, and propagation came from intertemporal substitution, capital accumulation, labor supply, and delayed capital installation. The striking claim was methodological as much as substantive: a small calibrated artificial economy, with parameters chosen from steady-state restrictions and outside evidence, could reproduce several postwar U.S. business-cycle moments, especially the relative volatility of consumption and investment, the comovement of major aggregates with output, and output persistence. Recessions were modeled as equilibrium responses to adverse real shocks in a neoclassical economy, rather than as failures of aggregate demand, price adjustment, or monetary coordination.

Long and Plosser (1983) supplied a complementary RBC benchmark by showing how sectoral productivity shocks could propagate through input–output linkages in a frictionless competitive economy. The key mechanism was a dated-input structure: goods produced in one sector become inputs for other sectors, and production decisions made today affect output only with a lag. A shock in one sector therefore changes current relative prices and output, and also changes future input supplies in other sectors. The model showed that persistence and comovement could arise from equilibrium responses to real sectoral shocks, without money, nominal rigidities, policy shocks, or disequilibrium adjustment. Relative to Kydland and Prescott (1982), propagation came not from delayed capital formation in one aggregate economy, but from production links across sectors.

A central quantitative weakness of early RBC models was that they generated too little volatility in total hours. In the standard divisible-labor setup, aggregate hours moved because employed workers changed hours per worker, the intensive margin, which required an implausibly large intertemporal labor-supply response. Kydland and Prescott (1982) tried to amplify this response through nonseparable leisure preferences, but hours still varied too little. G. D. Hansen (1985) addressed the problem by introducing indivisible labor: workers either work a fixed shift or do not work, so aggregate hours vary mainly through employment, the extensive margin, rather than through hours per worker. This allowed aggregate hours to respond more strongly to technology shocks without assuming a highly elastic individual labor supply. The cost was interpretive: the model matched employment fluctuations by treating nonemployment as part of an efficient equilibrium allocation, rather than as involuntary job loss caused by rationing, search frictions, or weak labor demand.

Prescott (1986) made the methodological claim explicit: growth and business cycles should be studied in the same stochastic growth model, with trend a measurement device rather than a distinct mechanism. King, Plosser and Rebelo (1988b, 1988a) made this operational by writing an RBC model with labor-augmenting technical change in stationary form. Starting from $Y_t = F(K_t, X_t N_t)$, dividing growing variables by X_t (so $y_t = Y_t/X_t$) yields stationary cyclical variables that are equilibrium objects derived from the model's own growth structure rather than residuals from an external trend.

The mature RBC program also developed a distinctive empirical discipline. Long and Plosser (1983) framed the benchmark question: what can a frictionless dynamic equilibrium economy explain before adding monetary shocks, nominal rigidities, policy mistakes, or market failures? Kydland and Prescott (1990) turned this benchmark into a program for reporting business-cycle facts. Theory was to guide which facts mattered: the relevant objects were detrended variables central to the growth model – output, consumption, investment, hours, productivity, wages, inventories, money, and prices, and the main moments were volatility, comovement with output, and lead–lag patterns. The resulting facts became targets for RBC models. Consumption was smoother than output, while investment and consumer durables were much more volatile. Labor input and real wages were procyclical, while capital and inventories adjusted more slowly. The nominal facts were less friendly to simple monetary stories: the price level was countercyclical, and neither the monetary base nor $M1$ systematically led output, while broader aggregates such as $M2$, especially $M2 - M1$, did lead output. This pointed toward credit arrangements as a possible extension of the baseline growth model. Hodrick and Prescott (1997) supplied the filtering procedure used to construct comparable cycle components in model and data.

Kydland and Prescott (1996) defended calibration as a computational experiment that is econometric in the broader sense, since it uses theory and measurement to derive a model economy's quantitative implications: theory specifies the primitives, equilibrium concept, and question while measurement disciplines the parameters, variables, and moments used to compare model with data. Calibration therefore became the construction of a model economy whose parameters were chosen from selected facts, such as factor shares, so that the solved model could answer a specific quantitative question, rather than simply a search for the best-fitting parameter vector. Model evaluation then followed a clear sequence: solve the model, simulate it, transform simulated and actual data in the same way, and compare the resulting moments. Discrepancies showed where measurement was weak, where parameters or shocks needed better evidence, or where the theory had to be extended. In this sense, RBC econometrics was a disciplined interaction between artificial economies and measured facts, not the estimation of autonomous behavioral equations. It completed the shift from Keynesian macroeconometric control systems to calibrated equilibrium laboratories and set the template for later DSGE work.

22.4.3 Empirical Tensions and Limitations of Early RBC Models

The calibration program gave early RBC models a clear empirical discipline, but it also made their substantive burdens visible. Once models were judged by simulated moments, the issue was no longer internal coherence alone, but whether the assumed shocks, labor-supply margins, and market structure were empirically credible. The early critiques by Summers (1986) and Mankiw (1989) therefore challenged the interpretation of the baseline RBC model rather than dynamic equilibrium modeling itself. Summers argued that moment matching was persuasive only if the parameters were independently justified, the Solow residual really measured exogenous technology, and the model could also confront prices, wages, interest rates, asset returns, and the apparent exchange failures of recessions, especially involuntary unemployment. Mankiw made the labor-market problem sharper. In a frictionless representative-agent model, a recession with falling consumption and employment must be explained as a voluntary decision to work less, which requires a large fall in the return to work and therefore places heavy weight on negative productivity shocks. The Solow residual therefore carried too much of the explanation. If measured productivity also reflects utilization, effort, labor hoarding, markups, or demand-driven changes in organization, it cannot be read mechanically as an exogenous technology shock. The first-generation critique was therefore not that RBC lacked theory, but that it asked too much of a narrow set of mechanisms: technology shocks, intertemporal substitution, flexible prices, and complete markets.

The subsequent literature developed these objections along four main margins. The first was labor-market comovement: the baseline model tied hours too closely to productivity and real wages, while the data showed a much weaker relation between hours and the return to work. The second was measurement: the volatility, persistence, and comovement used to evaluate RBC models often depended on the filter used to separate trend from cycle. The third was shock interpretation: the Solow residual was treated as technology, but it could also reflect utilization, effort, labor hoarding, markups, aggregation, and measurement error. The fourth was inference: calibration used external evidence to choose parameters, but often treated them as fixed values and judged fit by informal moment comparisons rather than explicit statistical criteria. Together, these criticisms shifted the debate from whether equilibrium models could be built and simulated to whether the benchmark RBC model had identified the right shocks, labor-market margins, moments, and standards of evidence.

Labor-market dynamics were the first empirical tension. In a technology-shock RBC model, higher productivity raises labor demand and the return to work, so hours, labor productivity, and real wages tend to move together. Postwar data showed the tension: hours were highly cyclical, but their comovement with productivity and real wages was weak. G. D. Hansen (1985) addressed the volatility problem by introducing indivisible labor and employment lotteries. G. D. Hansen and Wright (1992) showed that this was still insufficient: the benchmark model missed both the volatility of hours relative to productivity and the near-zero correlation between hours and productivity. Christiano and Eichenbaum (1992) traced the deeper problem to the one-shock structure: additional disturbances such as government-consumption shocks can move labor supply independently of measured productivity. The labor-market critique was therefore that technology shocks and intertemporal substitution alone could not explain the observed behavior of hours, productivity, and real wages.

The second tension was measurement. Hodrick and Prescott (1997) decomposed aggregates into a smooth trend τ_t and cycle $\tilde{y}_t = y_t - \tau_t$ and compared cyclical moments – a standardization that also made the empirical target depend on the trend–cycle choice. Harvey and Jaeger (1993b), Cogley and Nason (1995), and Canova (1998) showed that volatility, persistence, and comovement can change with the detrending method, so moments from the Hodrick–Prescott (HP) filter are partly filter-dependent. The criticism was therefore not simply that HP filtering was wrong, but that some RBC facts were partly filter-dependent. A model that matches HP-filtered moments may therefore be matching the measurement procedure as much as a robust property of the economy, so its conclusions should be checked against alternative trend–cycle decompositions.

The third empirical tension concerned the technology shock itself. In the baseline RBC model, technology shocks are not observed directly. They are inferred from the Solow residual and then used as the main stochastic impulse in the artificial economy. This move was central to Prescott (1986): it gave the model a measurable real shock, but it also made the theory vulnerable to a measurement objection. The residual measures technology cleanly only under restrictive assumptions: competitive factor pricing, constant returns, correctly measured factor services, and no omitted cyclical input margins. If actual production also depends on utilization and effort, then measured productivity may rise or fall because inputs are used more intensively, not because technology has changed. Summers (1986) and Mankiw (1989) argued that this was a central weakness of the early RBC interpretation: recessions could look like episodes in which technology went backward because the residual also absorbed utilization, effort, labor hoarding, markups, aggregation error, and demand-driven organizational changes.

Later work turned the measurement problem into an identification problem. Galí (1999) and Basu, Fernald and Kimball (2006) found that positive technology shocks reduce hours or inputs on impact, contrary to the baseline RBC prediction. The pattern is hard to reconcile with a flexible-price RBC mechanism but natural in sticky-price models, pointing toward nominal rigidities, utilization margins, or labor hoarding.

The fourth empirical tension concerned inference. Early RBC practice judged fit by informal comparison of selected moments with parameters treated as known. Gregory and Smith (1991) made the comparison statistical by treating the calibrated model as a null hypothesis with simulation-based moment distributions, and Canova (1994) extended this by treating the calibrated parameters themselves as uncertain. The econometric lesson was that calibrated RBC models needed explicit inference:

parameter uncertainty, sampling distributions for model moments, and a statistical distance between model and data.

The durable part of the equilibrium turn was methodological. Lucas (1976) made regime invariance central to policy evaluation: coefficients estimated under one policy rule cannot simply be carried into another. Rational expectations became a benchmark not because agents are always fully informed, but because systematic policy changes alter private forecasts and decision rules. RBC theory extended this logic into quantitative macroeconomics by requiring explicit stochastic equilibrium models: specify primitives and shocks, solve agents' decision rules, simulate the artificial economy, and compare its moments with the data. Later models often rejected the early RBC interpretation of recessions, but they retained its standards of policy-invariant primitives, solved decision rules, explicit shock processes, and systematic model–data comparison.

The less durable claims were substantive: policy-ineffectiveness results depended on restrictive information and nominal-adjustment assumptions that nominal rigidities later overturned, and early RBC models asked too much of technology shocks, flexible prices, and complete markets. The empirical critiques showed Solow residuals are not clean technology measures, filtered facts are sensitive to detrending, and calibration needs explicit inference, so equilibrium structure is necessary but not sufficient: credible quantitative models must also identify shocks, specify frictions, justify measurement, and report uncertainty.

The response split into two connected lines. One enriched equilibrium models with nominal rigidities, policy rules, adjustment costs, financial frictions, and statistically disciplined estimation, eventually leading to DSGE models. The other, discussed next, turned to vector autoregressions (VARs) and related time-series methods to describe aggregate dynamics and identify shocks without immediately imposing the full cross-equation restrictions of a calibrated equilibrium model. VAR evidence did not replace structural modeling, but it supplied dynamic facts, impulse responses, and measurement benchmarks that later structural models had to confront.

22.5 Re-Centering on the Data: The Reduced-Form Revolution and Cycle Measurement (late 1970s–1990s)

By the late 1970s, the structural interpretation of large-scale Cowles–Keynesian macroeconometric systems had become increasingly difficult to defend. The problem was not that these models were useless for forecasting or counterfactuals, but that their behavioral interpretation rested on exclusion restrictions, exogeneity assumptions, lag structures, and expectations proxies that were hard to defend as structural. These restrictions were not needed to describe dynamics, and they were not innocuous once the model was used for causal or policy interpretation. Lucas (1976) challenged the policy invariance of estimated behavioral equations. Sims (1980) challenged the identifying restrictions used to recover those equations in the first place. Together, the two critiques separated dynamic description from causal interpretation. Rather than start with a simultaneous-equation system whose equations already carried behavioral labels, the econometrician would first estimate the joint stochastic process of the observable variables and only then ask what additional assumptions were needed for causal interpretation.

The response was to retreat from *incredible* restrictions and re-center empirical macroeconomics on the joint stochastic dynamics of observable variables. In Sims's VAR program, the first step was to estimate the multivariate time-series process, summarize its dynamics through impulse responses and forecast-error variance decompositions, and impose structure only when a causal question required it. The reduced-form revolution therefore did not solve identification, but rather it changed the order of analysis. Dynamics came first, structural interpretation second. The following subsections trace how this logic shaped agnostic VARs, structural VARs, forecasting methods, and the measurement of trends and cycles.

22.5.1 From Structural Systems to Agnostic Dynamics

Sims (1972) provided an early test of the exogeneity assumptions built into macroeconometric models. Money was often placed in an exogenous policy block, so distributed-lag regressions of income on money were read as evidence about monetary causation. Sims asked whether that ordering had testable implications in the data. Using Granger's criterion – that X Granger-causes Y if lagged X improves forecasts of Y – he found evidence consistent with money helping to predict income. The main point was methodological rather than simply monetarist: exogeneity assumptions and zero restrictions should be checked against the dynamic predictive relations in the data before being used for structural interpretation.

Sargent and Sims (1977) provided a bridge between Cowles–Keynesian structural systems and the later VAR program. They represented business-cycle comovement with low-dimensional index models, in which many observable variables load on a few common stochastic components. Output, employment, consumption, and investment, for example, could share a common real-activity factor. The aim was not to recover deep structural shocks, but to summarize aggregate comovement with limited *a priori* theory. This shifted attention from pre-labeled structural equations to the joint stochastic representation of the data, preparing the ground for Sims's unrestricted VARs.

Sims (1980) made the reduced-form program operational through VARs. A VAR models each variable as a function of its own past values and the past values of the other variables in the system. Because each equation uses the same lagged information, the model can be estimated equation by equation using ordinary least squares. The approach was 'unrestricted' in the sense that it did not impose theory-based zero restrictions or assign behavioral labels to individual equations, although researchers still had to choose the variables, transformations, sample period, deterministic terms, and lag length. Once estimated, a stable VAR can be used to study how forecast errors propagate through the system over time. These propagation patterns are summarized by impulse responses, which trace the dynamic response of each variable to an innovation in the system.

The reduced-form innovations in a VAR, however, are not automatically structural economic shocks. They are forecast errors, and they are often contemporaneously correlated with one another. To interpret them as distinct economic shocks, the researcher must add an identifying assumption. One common early approach was recursive, or Cholesky, identification, which orders the variables and assumes that some variables can respond contemporaneously to others while others respond only with a lag. Impulse responses and forecast-error variance decompositions are then computed using these orthogonalized innovations. The key contribution of the VAR program was therefore to separate the estimation of reduced-form dynamics from the structural interpretation of shocks, making the identifying assumptions explicit rather than embedding them inside a large structural model.

Cooley and LeRoy (1985) gave the central critique of the VAR program's claim to theoretical neutrality: VARs did not eliminate identification, but relocated it. Cholesky orthogonalization, for example, converts reduced-form innovations into candidate shocks only by imposing a recursive contemporaneous ordering, which is itself an identifying restriction. Similarly, Granger-causality tests can reveal predictive precedence and lag exclusions, but they do not establish economic causality or policy exogeneity. The relevant distinction is between deep parameters that may support policy counterfactuals and shallow reduced-form objects that summarize a historical stochastic process, such as VAR coefficients, residual covariances ,orthogonalized impulse responses, and forecast-error variance decompositions. Because these shallow objects can change when policy rules change, they cannot by themselves support conditional policy experiments. This critique set the stage for the move to structural VARs, in which identifying assumptions are explicitly imposed rather than hidden in the statistical representation.

22.5.2 Structural VARs and the Search for Interpretation

Reduced-form VARs provide a disciplined way to describe how macroeconomic variables move together over time, but their forecast errors do not automatically have economic meanings. A VAR residual may combine several contemporaneous disturbances, so it can be interpreted as a monetary, technology, demand, or supply shock only after the econometrician imposes an identifying assumption. The problem is that the covariance patterns in the VAR residuals do not uniquely reveal the underlying structural shocks. Many different mappings from structural shocks to reduced-form forecast errors can produce the same observed covariance matrix. SVARs therefore make identification an explicit modeling choice: the researcher must state the assumptions that turn reduced-form innovations into economically interpretable shocks.

A first route replaced recursive Cholesky orderings with explicit contemporaneous restrictions. Bernanke (1986) modeled contemporaneous relations among VAR residuals to compare money-view, credit-view, and passive-money explanations of the money–income correlation, making orthogonalization a structural hypothesis. Blanchard and Watson (1986) applied the same impulse–propagation logic to business cycles, decomposing fluctuations into supply, demand, fiscal, and monetary disturbances. Together, these papers used VAR orthogonalization as economic identification.

Monetary-policy applications made short-run SVAR restrictions central. Researchers identified a monetary-policy shock by specifying what the central bank observes within the period and which private variables can respond immediately to policy. Sims (1992) showed both the power and the fragility of this strategy. Interest-rate innovations predicted output declines, but they also produced the price puzzle: prices sometimes rose after an identified monetary tightening. Sims read the puzzle as evidence that small VARs omitted information policymakers used when forecasting inflation, such as forward-looking price indicators. Christiano, Eichenbaum and Evans (2005) then used those impulse responses to estimate and evaluate a medium-scale DSGE model. This turned SVAR impulse responses into a bridge between reduced-form time-series evidence and structural model evaluation.

A second route identified shocks by restricting their long-run effects rather than their contemporaneous timing. Blanchard and Quah (1989), for example, separated demand shocks from supply shocks by assuming that demand shocks have no permanent effect on output. King, Plosser, Stock and Watson (1991) extended this long-run approach to cointegrated systems with balanced-growth restrictions. Galí (1999) then used a long-run restriction to identify technology shocks and challenged the RBC prediction that a positive technology shock should raise productivity, output, and hours together.

A third route relaxed the reliance on exact zero restrictions through sign restrictions. Instead of requiring some responses to be exactly zero, this approach identifies shocks by asking whether their effects match theoretically predicted signs. Canova and De Nicoló (2002) used this idea to assign economic meaning to orthogonalized innovations based on the co-movement patterns they generated. Uhlig (2005) applied the same logic to monetary policy, asking what happens to output after a shock that looks like a monetary tightening in nominal and reserve-market variables. He restricted the responses of the federal funds rate, prices, and nonborrowed reserves, but deliberately imposed no sign restriction on real gross domestic product (GDP), so that an output decline was not built into the identification scheme. These methods reduced reliance on exact timing and long-run zero restrictions, but changed the object of inference. The researcher no longer obtained a single structural decomposition. Instead, the data and restrictions defined a set of admissible impulse responses.

The later debate clarified the limits of the SVAR strategy. Chari, Kehoe and McGrattan (2008) showed that long-run SVAR identification of technology shocks suffers from finite-sample lag-truncation bias when nontechnology shocks matter. Christiano, Eichenbaum and Vigfusson (2006) defended SVARs as model-evaluation tools under explicit invertibility and information conditions, with short-run restrictions more reliable than long-run ones. SVARs, therefore, turn identification into an explicit modeling choice. Each identification family carries a different vulnerability: short-run restrictions rely on timing and information assumptions; long-run restrictions rely on low-frequency implications and imprecise estimates of cumulative effects; sign restrictions trade exact zeros for a set of admissible impulse responses. These vulnerabilities paved the way to larger information sets, Bayesian shrinkage, and factor-augmented VARs.

22.5.3 High-Dimensional and Bayesian VARs in Forecasting

The reduced-form VAR program also faced a practical dimensionality problem. A VAR with many variables and lags requires estimating a large number of coefficients, so expanding the system can quickly increase estimation uncertainty and weaken out-of-sample forecasting performance. Small VARs were easier to estimate, but they often contained too little information to match the information sets used by forecasters, policymakers, and financial markets. The price puzzle illustrated this limitation: a low-dimensional monetary VAR could misidentify policy shocks if it omitted forward-looking indicators observed by the central bank. Two main responses emerged. Bayesian VARs retained the VAR framework but used shrinkage to limit overfitting, while factor models summarized large data panels with a small number of estimated common components.

Doan, Litterman and Sims (1984) and Litterman (1986) turned Bayesian shrinkage into the first practical answer to VAR overparameterization. Their approach became known as the Minnesota prior. The basic idea was to start from a simple forecasting benchmark: each variable is predicted mainly by its own recent past, often in a highly persistent or near-random-walk form. The prior then shrinks the VAR toward this benchmark while still allowing the data to estimate departures from it. In practice, it gives more flexibility to a variable's own recent lags and imposes tighter shrinkage on distant lags and lags of other variables. This reduces estimation noise without imposing the zero restrictions associated with Cowles-style structural models.

This made medium-sized VARs useful for forecasting, predictive distributions, and conditional projections. The contribution remained deliberately reduced-form: BVARs characterized dynamic statistical interdependence and produced reproducible forecasts, while causal policy interpretation still required additional identifying assumptions. Litterman's five-year record at the Minneapolis Fed showed that a small BVAR could match the average accuracy of commercial forecasting services while remaining inexpensive, reproducible, and free of judgmental adjustment. This made Bayesian shrinkage a durable forecasting technology.

Later work turned BVARs into a broader Bayesian forecasting toolkit. Kadiyala and Karlsson (1997) made richer priors feasible through posterior simulation (Gibbs sampling), Banbura, Giannone and Reichlin (2010) showed that large BVARs could handle a hundred or more variables when prior tightness scales with system size, and Giannone, Lenza and Primiceri (2015) reduced researcher discretion by selecting prior tightness hierarchically from the data. By the 2010s BVARs were standard in institutional forecasting toolkits.

The factor-model strand offered a different solution. Instead of estimating a large VAR in all observed variables, factor methods assumed that many macroeconomic series share a small number of common forces. Stock and Watson (1989) modernized the Burns–Mitchell coincident-indicator tradition in a dynamic-factor framework. Stock and Watson (2002) then turned factor extraction into a forecasting method: principal components summarized hundreds of monthly predictors into a few diffusion indexes, which improved forecasts of real activity and inflation relative to univariate models, small VARs, and leading-indicator benchmarks. For forecasting, the factors did not need unique structural names; they only had to summarize predictive information.

Bernanke, Boivin and Eliasz (2005) connected factor methods back to SVAR identification through the factor-augmented VAR, or FAVAR. Standard monetary VARs used sparse information sets, while central banks and private agents observed far more data. A FAVAR extracted latent factors from a large panel and estimated monetary policy jointly with those factors. This richer information set helped separate true policy innovations from systematic policy responses to information omitted from a small VAR. It also allowed the researcher to trace monetary-policy effects on many more variables than a standard VAR could include. FAVARs therefore addressed both problems created by small monetary VARs: contaminated policy shocks and narrow impulse-response evidence. Their limitation was that factor augmentation improved the information set but did not eliminate the need for structural identification.

A final extension relaxed the assumption of constant dynamics. Cogley and Sargent (2005) and Primiceri (2005) introduced time-varying coefficients and stochastic volatility into VAR analysis, allowing both shock variances and propagation mechanisms to change over time. These models became central to the Great Moderation debate. Cogley and Sargent found evidence of drifting inflation persistence and changing monetary-policy behavior, while Primiceri concluded that smaller non-policy disturbances explained much of the post-1980 decline in volatility. Time-varying-parameter (TVP) VARs made instability empirically visible, but their flexibility came with a cost: results depended on priors governing the speed of coefficient drift and volatility changes.

The lasting lesson is practical. High-dimensional VAR research made reduced-form forecasting scalable without restoring structural certainty: Bayesian shrinkage handles medium and large VARs, factor models compress very large panels, FAVARs improve policy-shock measurement, and TVP-VARs allow the law of motion to evolve. In the mature view of Stock and Watson (2006) and Karlsson (2013), BVARs and factor models are complements that deliver transparent, reproducible forecasts without the full structural commitment of large macroeconometric models.

22.5.4 Stochastic Trends and Cycle Measurement

By the early 1980s, business-cycle measurement faced a basic problem: the trend itself might move stochastically. The older deterministic-detrending approach treated output as a smooth growth path plus a stationary cycle. Under that view, shocks move the economy away from trend only temporarily. The stochastic-trend literature changed the object of measurement. If output is integrated, shocks can change the long-run forecast of the series, so the boundary between trend and cycle depends on the assumed data-generating process rather than on a mechanical detrending rule.

Beveridge and Nelson (1981) gave stochastic trends their foundational model-based interpretation. Instead of treating trend as a smooth curve fitted through the data, they defined it as the long-run forecast of the series based on information available at date t, after adjusting for average growth. The observed series is decomposed as $y_t = T_t + c_t$, where T_t is the permanent component and c_t is the transitory, or cyclical, component. The permanent component follows a random walk with drift: when new information changes the forecast of the series far into the future, the trend itself moves. The cyclical component is the remaining temporary deviation from that forecast. The strength of the Beveridge–Nelson decomposition is that it defines trend from the estimated time-series model rather than from an external smoothing rule, and it can be computed using only past and current data. Its weakness is that the measured cycle depends on the autoregressive integrated moving average (ARIMA) model chosen for the growth rate of the series.

Nelson and Plosser (1982) turned this issue into a major empirical challenge to deterministic detrending. Using long U.S. macroeconomic series, they found that unit-root tests often failed to reject stochastic trends. This weakened the practice of fitting a linear or quadratic trend and treating the residual as the business cycle. It also supported the emerging view that real shocks could account for a substantial part of output fluctuations. Cochrane (1988) then sharpened the question. The issue was not only whether output contained a unit root, but how large the permanent component was. His variance-ratio approach compared the variance of long differences with the variance implied by a pure random walk. For U.S. gross national product (GNP), the random-walk component appeared smaller than the strongest reading of Nelson–Plosser suggested, leaving room for both permanent and transitory movements.

The next step showed that stochastic detrending itself required identification. Watson (1986) compared Beveridge–Nelson and unobserved-components decompositions, where output is written as the sum of a stochastic trend and a stationary cycle. In principle, ARIMA and unobserved-components models can describe the same observed series. In practice, they can fit short-run data similarly while implying very different long-run responses to shocks. Morley, Nelson and Zivot (2003) later explained why the decompositions often differed so sharply. Standard unobserved-components models usually set the covariance between trend and cycle innovations to zero. Once that restriction is relaxed, the Beveridge–Nelson and unobserved-components decompositions can coincide, and the zero-covariance restriction is rejected for U.S. GDP. The lesson is that permanent–transitory decompositions are model-based objects, not facts read directly from the data.

The unit-root insight also influenced multivariate business-cycle measurement. Stock and Watson (1988) asked whether integrated series share common stochastic trends, so that some linear combinations are stationary. King et al. (1991) connected this logic to balanced growth and common-trend restrictions in systems containing output, consumption, investment, money, and prices. The trend-cycle problem therefore moved from a univariate question: how much of output is permanent, to a system question: how many stochastic trends drive the economy, and how do variables adjust around them?

A parallel branch returned to the older NBER concern with expansions and contractions. Hamilton (1989) modeled real GNP growth as switching between latent regimes governed by a Markov process. This approach treated recessions as probabilistic states rather than as residuals from a trend.

The lasting lesson was that cycle measurement is never neutral. Beveridge–Nelson decompositions define the permanent component through long-run forecasts; unobserved-components models impose assumptions about trend and cycle innovations; common-trend models impose restrictions across variables; Markov-switching models define cycles as latent regimes. Each method answers a different measurement question. Practitioners still wanted a simpler, transparent device for comparing actual and simulated series, and that demand led to the widespread use of the HP filter discussed next.

22.5.5 Statistical Filters and Their Limitations

Model-based decompositions such as Beveridge–Nelson and unobserved-components models required the analyst to specify a stochastic process for trend and cycle. A different strand chose a simpler route: define the cycle by applying a statistical filter directly to the data. These filters made business-cycle measurement reproducible and easy to compare across series, but also redifining the cycle as the component extracted by a particular transformation, not an observed economic object.

The most influential statistical filter was the HP filter, proposed by Hodrick and Prescott (1997) and circulated earlier as a 1981 working paper. The HP filter separates a log series into a smooth trend and a cyclical component, defined as the gap between the observed series and that trend. The trend is chosen by balancing two objectives: it should remain reasonably close to the data, while its growth rate should change only gradually. The smoothing parameter controls this tradeoff. A smaller smoothing parameter allows the trend to follow the data more closely, leaving less movement in the measured cycle. A larger parameter forces a smoother trend, leaving more short-run variation in the cyclical component. The filter became attractive because it gave researchers a common and easily reproducible way to transform actual and simulated series before comparing volatilities, correlations, and lead–lag patterns. Its weakness followed from the same feature: it defined the business cycle by a statistical smoothness rule rather than by an economic model of potential output or cyclical adjustment.

King and Rebelo (1993) made this point precise by analyzing the HP filter in the frequency domain. The filter removes low-frequency movements and keeps higher-frequency deviations, but it can also change measured persistence, volatility, and comovement. Harvey and Jaeger (1993a) pushed the criticism further from a structural time-series perspective: mechanical detrending can lead researchers to report spurious cyclical behavior, and stylized facts should be consistent with the stochastic properties of the data. Canova (1998) generalized the robustness problem by showing that business-cycle facts can change, sometimes qualitatively, across detrending methods. The criticism was therefore not simply that HP filtering was wrong. It was that filtered moments were conditional on the filter.

Band-pass filters offered a more explicit alternative. Instead of choosing a smoothness parameter, Baxter and King (1999) defined the business cycle as fluctuations with periods between 6 and 32 quarters and constructed a moving-average filter that passes that frequency band while suppressing lower-frequency trend movements and higher-frequency noise. Stock and Watson (1999) used this frequency-domain approach to summarize postwar U.S. business-cycle comovement across a broad set of macroeconomic series. These methods clarified the measurement choice, but did not remove it: the analyst still had to decide which frequencies count as the business cycle.

The spurious-cycle critique showed why this mattered: Cogley and Nason (1995), Murray (2003), and Phillips and Jin (2021) showed that HP and band-pass filtering of highly persistent or integrated series can generate cycle-like residuals even when the underlying process contains no genuine business-cycle propagation.

A separate problem appeared in real time. Orphanides and van Norden (2002) showed that output-gap estimates are highly unreliable at the end of the sample, where policymakers need them. Revisions came partly from revised data, but mainly from the difficulty of estimating the current trend without future observations; output gaps enter monetary-policy rules, fiscal assessments, and inflation forecasts, so real-time trend errors can become policy errors.

Hamilton (2018) gave the most forceful modern critique of HP filtering and proposed a forecasting-based alternative. Instead of drawing a smooth trend through the whole sample, Hamilton asks what recent values of the series would have predicted eight quarters ahead. For quarterly data, he recommends forecasting y_{t+8} from the four most recent observations available at date t. The cyclical component is the forecast error: the part of y_{t+8} that recent history failed to predict. This method uses only current and past data, avoids the two-sided smoothing and endpoint problems of the HP filter, and needs no smoothing parameter. It still requires choices about the forecast horizon and lag length, but it defines the cycle as a medium-run forecast error rather than as a residual from a smooth trend.

Statistical filters remain valuable descriptive tools, especially when the goal is to compare many series or to transform simulated and actual data in the same way. Filters define the cycle they measure. HP filtering defines it through smoothness, band-pass filtering through frequency, unobserved-components models through a stochastic decomposition, and Hamilton's method through forecast errors. Business-cycle facts that survive several reasonable transformations are more credible than facts that appear only under one filter. Statistical filtering therefore completed the reduced-form turn in cycle measurement, but also revealed its limit: even minimal measurement procedures carry identifying assumptions.

The reduced-form revolution aged best as a tool of empirical description. It forced macroeconomists to take the joint dynamics of the data seriously, to report impulse responses and forecast errors transparently, to confront the information limits of small systems, and to treat forecasting performance as an empirical benchmark rather than a secondary task. VARs, Bayesian shrinkage, factor models, and statistical decompositions all strengthened macroeconometrics by making dynamics, uncertainty, and measurement explicit. What aged less well was the claim that these methods could remain theoretically neutral. VAR innovations do not carry economic labels by themselves. Structural interpretation requires identifying assumptions. Statistical filters do not reveal a pre-existing cycle, but rather they define one through assumptions about smoothness, frequency, stochastic trends, or forecast errors. The lasting lesson is therefore balanced: reduced-form methods remain indispensable benchmarks for describing macroeconomic dynamics, forecasting, and organizing facts, but they cannot by themselves deliver causality, policy counterfactuals, or model-free measures of trend and cycle. The next step in the literature was to reintroduce structural content more selectively, through external instruments, narrative and high-frequency identification, local projections, state-space methods, and real-time measurement.

22.6 Identification and Measurement After VARs (1990s–2000s)

By the late 1990s, the central problem in empirical business-cycle analysis was no longer whether VARs could describe macroeconomic dynamics, but whether their internal restrictions could support causal interpretation. Recursive orderings, long-run zero restrictions, and sign restrictions made identification assumptions explicit, but the credibility of a monetary, fiscal, oil, or technology shock still depended on timing, neutrality, or sign claims imposed inside the system. The post-1990s response was to bring more information into both identification and measurement. Narrative methods used policy records, legislation, military events, and oil disruptions; high-frequency identification used asset-price movements around policy announcements; proxy SVARs and local projections used those external series to estimate dynamic responses; and state-space, mixed-frequency, and real-time methods treated latent variables, data vintages, and release calendars as objects to be modeled directly. The unifying move was not to escape identification, but to place it on more transparent ground: shocks, gaps, trends, regimes, and business conditions had to be tied to observable institutions, market reactions, data releases, or explicit latent-state models.

22.6.1 Narrative and High-Frequency Identification

By the late 1990s, the main weakness of structural VAR identification had become clear: many policy shocks were identified by assumptions internal to the VAR. Recursive orderings required timing assumptions, long-run restrictions required neutrality assumptions, and sign restrictions required theory-based response patterns. These assumptions made identification explicit, but they still left open the central question: did the estimated shock represent an exogenous policy or supply disturbance, or did it partly reflect the systematic response of policymakers and markets to information about the economy? The next response was to bring in information from outside the VAR. Two literatures developed in parallel. Narrative identification used historical records, policy documents, speeches, legislation, and event dates to isolate plausibly exogenous policy or supply movements. High-frequency identification used asset-price changes in narrow windows around policy announcements to measure the surprise component of those announcements.

The modern narrative approach begins with Romer and Romer (1989). They returned to the Friedman–Schwartz tradition but made it more systematic: read the historical record and identify episodes in which the Federal Reserve deliberately tightened policy to reduce inflation, rather than because real activity was already weakening or strengthening. The shock measure was therefore not a money aggregate or a VAR residual. It was an externally dated sequence of policy episodes based on Federal Open Market Committee (FOMC) records, speeches, and policy discussions. This gave the narrative approach its force: it tried to identify policy actions whose motivation could be defended historically. It also gave the approach its main limitation: the dating was judgmental and retrospective, so the classification of shocks had to remain open to debate.

Romer and Romer (2004) made the same logic quantitative by constructing an intended federal-funds-rate series from FOMC records and removing the component predictable from the Federal Reserve's Greenbook forecasts of output, inflation, and unemployment. The residual is the monetary shock – a measure that combines narrative evidence with the Fed's real-time information set and directly targets the main endogeneity problem in monetary VARs.

Narrative identification also reshaped fiscal policy measurement. Romer and Romer (2010) used presidential speeches, executive-branch documents, and congressional reports to classify major postwar tax changes by motivation. Tax changes taken for countercyclical reasons or tied to spending changes were treated differently from changes motivated by long-run growth or deficit reduction. The central idea was that a legislated tax change can be used to estimate fiscal effects only if its timing and motivation are plausibly unrelated to the short-run path of output. Ramey (2011) applied the same logic to government spending, especially military buildups. Ramey's key point was timing: private agents often learn about future defense spending before government purchases actually rise. If the shock is dated when spending appears in the data, a VAR may identify the impulse too late. Dating the shock when the news arrives changes the estimated responses of consumption, real wages, and output.

Hamilton (2003) extended the identification logic to oil shocks. Since oil-price changes confound supply disruptions, global-demand swings, and expectations about future scarcity, Hamilton emphasized historically grounded oil disruptions and nonlinear price movements (large increases relative to recent experience), moving the object of measurement from the raw oil-price innovation toward economically interpretable shocks.

High-frequency identification took a different route, using financial-market prices to measure policy news in short windows. Kuttner (2001) introduced the canonical monetary application using federal-funds futures to separate expected from unexpected policy changes. The idea is that changes in futures rates around an FOMC announcement reveal how markets revised their expectations of the federal funds rate. Because the futures contract reflects the average rate over the month, the change is adjusted for the share of the month remaining after the announcement. Kuttner showed that market interest rates responded strongly to the unexpected component of policy, but only weakly to the anticipated component.

Gürkaynak, Sack and Swanson (2005) showed that a single target-rate surprise could not explain the movement of asset prices around FOMC announcements. Their two-factor decomposition separates a *target factor* (news about the current funds-rate target) from a *path factor* (news about the expected future policy path, closely tied to forward guidance and especially relevant for longer Treasury yields). Monetary-policy announcements were no longer one-dimensional changes in the current short rate but also changed expectations about future policy.

Bernanke and Kuttner (2005) used the high-frequency surprise to study the stock market: an unanticipated easing raises broad stock indexes, and much of the response comes through expected excess returns rather than mechanically through the risk-free discount rate. High-frequency identification thus broadened monetary-transmission analysis from interest rates to asset prices, risk premia, and financial conditions.

The high-frequency literature then encountered its own identification problem. Nakamura and Steinsson (2018) showed that FOMC announcements can reveal information about the Fed's view of the economy as well as information about policy. Around announcements, a surprise increase in interest rates was associated with higher expected output growth, which is the opposite of what a pure contractionary monetary shock would predict. This is the *Fed information effect*: markets may infer that the central bank is tightening because it has good news about future economic fundamentals. Jarociński and Karadi (2020) separated these components using the joint movement of interest rates and stock prices, confirming that high-frequency windows improve but do not eliminate the interpretation problem.

Narrative and high-frequency identification therefore changed the source of identifying variation. Shocks no longer had to come from a recursive ordering, a long-run zero, or a sign pattern inside a VAR. They could be built from institutional records, fiscal legislation, military news, oil disruptions, and asset-price surprises around policy announcements. These methods made identification more concrete and more transparent, but they did not make it automatic. Narrative measures depend on classification and timing. High-frequency surprises depend on event windows and on the information content of announcements. The next subsection asks how these externally constructed shock measures can be used inside dynamic systems through external-instrument and proxy-SVAR methods.

22.6.2 External Instruments and Proxy SVARs

The narrative and high-frequency approaches produced external series z_t that contain information about policy or other shocks, such as the Romer–Romer monetary residual, the Romer–Romer tax series, and high-frequency monetary surprises. Proxy-SVAR methods treat these series as instruments for latent structural shocks inside a VAR, rather than as observed shocks themselves. This distinction matters because narrative series may contain dating, classification, or measurement error, and high-frequency surprises may mix pure policy news with central-bank information. In the VAR, the reduced-form residual u_t is a mixture of structural shocks. If z_t satisfies the usual IV relevance and exogeneity conditions for the target shock, its covariance with u_t picks out the part of the residual vector associated with that shock. The researcher can then trace that component through the VAR to estimate impulse responses. Identification therefore comes from the external proxy, not from a recursive ordering, a long-run zero restriction, or a sign pattern imposed inside the VAR.

Mertens and Ravn (2013, 2014) provided the leading fiscal application of the proxy-SVAR logic. Mertens and Ravn (2013) used narratively identified personal and corporate income tax changes as external proxies for structural tax shocks in a quarterly VAR. The narrative record identified tax changes that were plausibly unrelated to current macroeconomic conditions, while the VAR traced how output, employment, consumption, investment, and tax revenues responded over time. The paper also showed that tax policy should not be treated as one aggregate shock: personal and corporate tax changes work through different channels.

Mertens and Ravn (2014) then used the same approach to explain why previous estimates of tax multipliers disagreed. Standard SVARs often inferred tax shocks from tax-revenue movements, but tax revenues also move automatically when GDP changes. Narrative measures used legislative history, but they could be noisy. Treating the narrative series as an instrument inside the VAR combined the two approaches: the narrative evidence supplied the identifying information, and the VAR estimated the dynamic effects. The reconciliation was therefore straightforward: the disagreement came less from narrative versus SVAR methods as such, and more from how each method handled automatic revenue movements and measurement error.

Gertler and Karadi (2015) brought the same framework to monetary policy. They used high-frequency monetary surprises around FOMC announcements as external instruments for monetary-policy shocks in a VAR with macroeconomic and financial variables. The point was to trace how a monetary shock propagates through credit costs, term premia, credit spreads, output, and inflation. Their main substantive result was that modest movements in short rates can generate large movements in credit costs and real activity, mainly because term premia and credit spreads respond strongly. This shifted monetary transmission analysis toward financial frictions rather than a frictionless interest-rate channel alone.

Stock and Watson (2018) formalized external-instrument macroeconometrics, distinguishing internal instruments (restrictions inside the VAR) from external instruments (narrative, high-frequency, or institutional sources) and clarifying the difference between the two-step proxy-SVAR method (estimate the VAR, then identify the shock in its residuals) and the one-step IV approach (estimate dynamic causal effects directly, horizon by horizon). Proxy SVARs need the VAR information set; one-step IV can be more robust in some cases. Weak-instrument-robust inference Montiel Olea, Stock and Watson (2021) subsequently made instrument strength part of the reporting standard for proxy SVARs.

The proxy-SVAR program changed the source of identification but not the propagation device – effects are still traced through a fitted VAR, which leads to local projections, where the same external variation drives horizon-by-horizon regressions instead.

22.6.3 Local Projections

The proxy-SVAR approach identifies a shock with an external instrument and then uses a fitted VAR to trace its effects. Jordà (2005) proposed a more direct alternative: estimate the response separately at each horizon. A local projection asks how the outcome h periods later responds to the shock today, after controlling for predetermined variables such as lags of output, prices, interest rates, or other state variables. In its simplest form, the regression is $y_{t+h} = \alpha_h + \beta_h s_t + \gamma_h' x_t + u_{t+h}$. The coefficient β_h is the estimated response at horizon h. Repeating this regression for $h = 0, 1, \ldots, H$ gives the full impulse-response path. If s_t is an observed shock, the regression can be estimated directly; if the researcher only has an external proxy for the shock, the same equation can be estimated by IV. Local projections became popular because they are transparent, easy to combine with narrative or high-frequency shocks, and especially convenient for nonlinear or state-dependent responses, such as different effects in recessions and expansions. Their main cost is precision: because each horizon is estimated separately, the estimated response path can be noisy.

The first major policy applications used this flexibility to study state dependence. Auerbach and Gorodnichenko (2012) estimated fiscal multipliers that vary over the business cycle and found much larger government-spending effects in recessions than in expansions. Their contribution was to shift the question from ‘what is the fiscal multiplier?’ to ‘what is the fiscal multiplier in a particular state of the economy?’ These papers made state-dependent impulse responses central for post-crisis fiscal policy analysis.

Other applications used local-projection-style designs to exploit different sources of identifying variation. Nakamura and Steinsson (2014) estimated fiscal effects from cross-state variation in U.S. military procurement. Because U.S. states share the same monetary policy and federal tax system, their design estimates an open-economy relative multiplier: how much more output rises in a region that receives more federal spending than another region. A counter-finding for monetary policy is that it appears less powerful in recessions (Tenreyro & Thwaites, 2016). Ramey and Zubairy (2018) then pushed back on strong fiscal state-dependence claims using long U.S. historical data. They found spending multipliers mostly in the range of 0.6 to 1.0 and little robust evidence that multipliers are systematically larger in high-slack periods. Together, these papers showed that local projections make heterogeneity easy to estimate, but the results still depend on the shock measure, sample, state definition, and controls.

The later methodological literature clarified what local projections do and do not change. Plagborg-Møller and Wolf (2021) then showed that, in population and with unrestricted lag structures, local projections and VARs estimate the same impulse responses. The difference is not the causal object being estimated, but the finite-sample tradeoff. VARs impose more structure across horizons and can be more precise if correctly specified; local projections impose less structure and can be more robust, but often at the cost of higher variance.

Local projections did not replace identification; they changed how dynamic responses are estimated once the shock has been identified. Their value lies in flexibility: they accommodate external instruments, nonlinearities, state dependence, panels, and long historical datasets with relatively transparent regressions. Their weakness is that flexibility can produce imprecise or specification-sensitive results. The central task remains the same as in SVARs and proxy SVARs: identify the shock credibly, choose the information set carefully, and report uncertainty honestly.

22.6.4 State-Space Models

Narrative, HFI, proxy IV and linear projection use external information to identify shocks and estimate their dynamic effects. State-space models address a different problem. Many objects that organize business-cycle analysis are not observed directly: the state of economic activity, potential output, the output gap, recession probabilities, stochastic volatility, and the natural rate of interest. State-space methods estimate these hidden objects by specifying how they evolve over time and how observed data reveal them imperfectly.

Kalman (1960) supplied the basic machinery for estimating unobserved states from observed data. In a state-space model, the economy has hidden state variables, such as potential output, trend productivity, or latent shocks, that evolve over time but are not observed directly. The observed data provide noisy signals about those states. The Kalman filter proceeds recursively: it first predicts the current state using past information, then updates that prediction when new data arrive. The update depends on the forecast error in the observed data, with more weight placed on the new observation when it is relatively precise.

Stock and Watson (1991) applied this logic to business-cycle measurement by treating coincident indicators (industrial production, employment, income, sales) as noisy signals of a single latent activity factor, replacing a judgmental composite indicator with an estimated probability model. The same dynamic-factor logic later anchored high-dimensional forecasting, FAVARs, and real-time monitoring. The Markov-switching extension of Hamilton (1989) expanded the state-space program beyond linear-Gaussian models by treating the hidden state as a discrete regime indicator estimated probabilistically.

The spread of state-space methods also required better computation. Carter and Kohn (1994) provided the forward-filtering/backward-sampling algorithm used in many Bayesian state-space models. The forward pass applies the Kalman filter; the backward pass draws the entire latent state path jointly. This made it practical to estimate models in which the hidden state is a whole time series, such as trend output, time-varying coefficients, or latent factors. S. Kim, Shephard and Chib (1998); C.-J. Kim and Nelson (1999) made stochastic-volatility and regime-switching state-space methods operational, underlying many later TVP-VARs and stochastic-volatility (SV) VARs.

The most influential policy application was the natural-rate framework of Laubach and Williams (2003), which jointly estimates potential output, trend growth, and the natural rate of interest r^* – the Wicksellian benchmark consistent with output at potential and stable inflation, by Kalman filtering of output, inflation, and interest-rate data. The framework made r^* operational at central banks, but the estimates are imprecise and depend heavily on signal-to-noise assumptions about the permanent versus transitory share of variation.

Holston, Laubach and Williams (2017) extended the framework to the US, Canada, the UK, and the euro area, finding parallel declines in trend growth and natural rates across advanced economies and turning r^* from a US-specific estimate into an international macroeconomic object.

By the mid-2010s, state-space methods had become a standard empirical-macro tool (Stock & Watson, 2016). The lasting contribution was not a single model of the cycle but a disciplined way to estimate hidden economic states from noisy data, with the standing caveat that every latent-state estimate depends on the assumed law of motion, measurement equation, and signal-to-noise restrictions.

22.6.5 Mixed-Frequency and Real-Time Measurement

Much empirical macroeconomics is written as if the data arrive cleanly: quarterly observations are aligned, complete, and already revised. Operational policymaking works under different conditions. Data arrive at different frequencies, with monthly, weekly, daily, and quarterly indicators released on different dates. The latest dataset is usually ragged: some series are available through the current month, others only through earlier months. The data are also revised, so the numbers available to policymakers in real time may differ substantially from the final data used by researchers years later. This literature made those complications part of the measurement problem rather than treating them as nuisances.

Croushore and Stark (2001) supplied the infrastructure by constructing real-time macroeconomic datasets organized by vintage, so forecasts and policy rules can be evaluated with the information actually available when decisions were made. The policy importance of real-time data became clear in Orphanides (2001). Taylor-type rules usually prescribe the interest rate as a function of inflation and the output gap. But the current output gap is not observed – it must be estimated from incomplete and revisable data. Orphanides showed that policy recommendations based on real-time data can differ sharply from recommendations computed with final revised data. Orphanides and van Norden (2002) made the point still more forcefully for output-gap estimates. Real-time output gaps are often revised by amounts as large as the gap itself, especially near turning points. The main problem is not only that GDP is revised. It is that the current trend, and hence current potential output, is hardest to estimate at the end of the sample, exactly when policymakers need it.

A second problem is mixed frequency. Key indicators do not arrive at the same interval: GDP is quarterly, employment and industrial production are monthly, unemployment claims are weekly, and financial prices are daily. Mariano and Murasawa (2003) handled this problem in state-space form. They combined monthly indicators with quarterly real GDP by treating quarterly GDP as an aggregation of an unobserved monthly process. The Kalman filter then recovers a monthly coincident index while respecting the timing of quarterly GDP releases. This approach extends the coincident-indicator tradition into a mixed-frequency state-space framework.

Ghysels, Santa-Clara and Valkanov (2004) proposed a different route through mixed-data sampling, or MIDAS. MIDAS keeps the regression framework. A low-frequency variable, such as quarterly GDP growth, is predicted directly from many lags of a high-frequency variable, such as monthly or daily data. The difficulty is that many high-frequency lags would normally imply too many coefficients. MIDAS solves this by forcing those lag coefficients to follow a simple smooth pattern governed by only a few parameters. Relative to state-space methods, MIDAS is simpler and more regression-based; relative to ordinary aggregation, it preserves more of the timing information in high-frequency data.

The corresponding real-time forecasting problem became known as nowcasting: estimating the present, the very near future, and the recent past before official data are complete. Giannone, Reichlin and Small (2008) gave nowcasting its canonical macroeconomic form. They used a dynamic factor model to combine a large panel of monthly indicators with quarterly GDP and update the current-quarter GDP estimate as each release arrives. The key object is the data release as news: the difference between the actual release and what the model expected. Banbura, Giannone and Reichlin (2011); Bańbura, Giannone, Modugno and Reichlin (2013) then turned this into a practical workflow that reads the release calendar and decomposes forecast revisions into news contributions.

Aruoba, Diebold and Scotti (2009) pushed the same logic to higher frequency. Their business-conditions index combines daily, weekly, monthly, and quarterly indicators in a dynamic-factor state-space model. The index treats business conditions as a latent state inferred from noisy signals arriving at different frequencies. This makes it possible to update the assessment of current economic activity at daily frequency, rather than waiting for monthly or quarterly releases.

Wu and Xia (2016) applied related measurement logic to monetary policy at the zero lower bound. After 2008, the federal funds rate no longer summarized the stance of policy, because it was constrained near zero while the Federal Reserve used forward guidance and asset purchases to affect longer-term rates. Wu and Xia estimated a shadow federal funds rate from the yield curve. The shadow rate is the latent short rate that would prevail without the zero lower bound; it can turn negative even when the observed federal funds rate is stuck near zero. This made unconventional monetary policy comparable to conventional policy in a single interest-rate measure.

The lasting contribution was practical: the empirical object shifted from final, aligned, quarterly datasets to the information flow policymakers and markets actually observe. Real-time vintages, mixed-frequency state-space and MIDAS methods, nowcasting, and shadow-rate models all illustrate that measurement is not neutral – the answer depends on the data vintage, release calendar, frequency conversion, and latent variable being estimated.

The post-VAR identification and measurement agenda aged well as a discipline of transparency. It moved empirical macroeconomics away from treating VAR residuals, filtered gaps, or policy innovations as self-interpreting objects. Narrative records, high-frequency market reactions, external instruments, local projections, state-space models, and real-time datasets all made the source of information more explicit. A monetary shock could be tied to an FOMC announcement, a tax shock to legislative history, a fiscal shock to military news, an output gap to a stated filtering or state-space model, and a nowcast to the actual sequence of data releases. This did not make inference mechanical, but it made the identifying and measurement assumptions visible. That was the durable advance: empirical macroeconomics gained better shock measures, more transparent impulse responses, richer latent-state estimates, and tools that respect the information actually available to policymakers.

What aged less well was the hope that moving outside the VAR would settle causality once and for all. Narrative measures still require classification and timing judgments; high-frequency surprises can mix policy news with central-bank information; external instruments can be weak or invalid; local projections remain sensitive to controls, samples, and state definitions; and state-space estimates of gaps, regimes, volatility, and natural rates depend on signal-to-noise assumptions that are hard to verify. Internal VAR restrictions were not rendered useless, but their limits became clearer: credible business-cycle analysis requires evidence that survives alternative identifying assumptions, data vintages, information sets, and measurement models. The main legacy of this section is therefore a more modest and more useful empirical standard. Identification and measurement cannot be avoided. They must be stated, defended, tested where possible, and compared across methods. The next structural models inherited these objects – identified impulse responses, shock series, factor estimates, latent states, and real-time measures, as empirical targets rather than as facts requiring no interpretation.

22.7 Estimation and Synthesis in New Keynesian DSGE Models (1990s–2010s)

By the late 1990s, the structural ambition that the reduced-form revolution had bracketed returned in a new form. The New Keynesian DSGE program absorbed the RBC methodological discipline – optimizing agents, rational expectations, equilibrium decision rules, and explicit shock processes, while adding nominal rigidities, monetary-policy rules, and later real, financial, and labor-market frictions. Three developments made the synthesis operational: sticky-price and sticky-wage theory supplied the economic mechanism, rational-expectations solution methods and state-space likelihoods made the models estimable, and Bayesian inference disciplined weakly identified parameters. The result was the medium-scale DSGE paradigm of Christiano et al. (2005) and Smets and Wouters (2007), which the post-2007 crisis then stress-tested.

22.7.1 Nominal Rigidities and the Limits of RBC

RBC made macroeconomics dynamic, stochastic, and explicitly grounded in optimization, but by the 1990s the baseline mechanism faced three pressures: weak internal propagation (with persistence inherited from the assumed shock process), absent monetary non-neutrality despite VAR and narrative evidence to the contrary, and Galí (1999)'s long-run-identified finding that positive technology shocks lower hours on impact. The New Keynesian response kept optimizing agents, rational expectations, and general equilibrium, but added monopolistic competition and nominal rigidities so that demand, inflation, and monetary policy can matter in the short run.

The basic nominal-rigidity mechanisms came from three closely related approaches. Taylor (1980) introduced staggered nominal wage contracts whose offset reset dates make the aggregate wage adjust gradually even when newly negotiated wages respond to current conditions, creating persistence beyond the contract length. Rotemberg (1982) modeled price stickiness through quadratic costs of nominal price change, generating smooth rather than discrete adjustment. Calvo (1983) let only a random fraction $1-\theta$ of firms reset prices each period, so θ controls the average duration of price stickiness. Calvo pricing became dominant for analytical tractability. Rotemberg pricing remained useful because, under a suitable parameter mapping, it generates the same log-linear Phillips-curve slope without explicit price-vintage heterogeneity.

The integration into dynamic general equilibrium took place in the 1990s. Yun (1996) embedded Calvo pricing in a stochastic-growth model with monopolistic competition and monetary shocks, reproducing inflation–output comovement better than flexible-price models. Goodfriend and King (1997) named the resulting New Neoclassical Synthesis: RBC methods on the real side, Keynesian nominal frictions on the monetary side. Rotemberg and Woodford (1997) turned the framework into an optimization-based econometric model for evaluating policy rules: policy evaluation must use private decision rules derived from optimization rather than reduced-form correlations that shift with the rule. Erceg, Henderson and Levin (2000) extended the logic from prices to wages, so monetary policy faces a genuine tradeoff among stabilizing the output gap, price inflation, and wage inflation, and strict price-inflation targeting is no longer generally optimal.

Clarida, Galí and Gertler (1999) gave the small New Keynesian model its standard policy form: a forward-looking demand curve (the Wicksellian channel through which the expected real rate relative to the natural rate moves private spending), a New Keynesian Phillips curve in which inflation depends on expected future inflation and the output gap, and a Taylor-type rule that sets the nominal rate in response to inflation and the gap with optional smoothing. This compact three-block system became the benchmark for monetary-policy analysis and the starting point for the larger DSGE models developed later.

The framework also clarified what RBC had missed: in the flexible-price RBC model the natural allocation is usually the actual one and monetary policy has little systematic role, while in the NK model sticky prices and wages prevent the economy from reaching the flexible-price benchmark immediately, so monetary policy matters through its effect on real rates, aggregate demand, and the output gap — explaining temporary monetary non-neutrality without abandoning rational expectations.

The empirical credibility of this program depended on whether the assumed nominal rigidities resembled actual price-setting behavior. Bils and Klenow (2004) used Bureau of Labor Statistics (BLS) microdata and found that many consumer prices changed more frequently than standard calibrations had assumed, with large differences across categories. This raised a problem for simple Calvo calibrations: if prices change often, why do monetary shocks appear to have persistent real effects? Later micro-price evidence gave a more nuanced picture. Klenow and Kryvtsov (2008) attributed most variation in monthly inflation to the size, not the frequency, of price changes. Nakamura and Steinsson (2008) showed that regular prices are stickier once sales and product substitutions are treated carefully, but also documented facts that simple Calvo models do not capture well: many regular price changes are decreases, price changes are seasonal, and the probability of adjustment varies with price duration. The conclusion was not that nominal rigidity was empirically irrelevant. It was that a single Calvo parameter is best interpreted as a tractable aggregate shortcut, not a literal description of firm pricing.

By the mid-2000s, the New Keynesian synthesis had become the main policy-oriented successor to RBC (Mankiw, 2006; Galí & Gertler, 2007). What aged well was the synthesis itself: business-cycle models had to remain dynamic, stochastic, forward-looking, and structural, but they also had to explain inflation dynamics, monetary transmission, and inefficient fluctuations. What aged less well was the idea that the small three-equation model, with a representative household, a representative sticky-price sector, and a single Calvo parameter, could carry the full empirical burden. The next step was therefore to enrich the DSGE structure with additional real and nominal frictions, micro-founded pricing evidence, and likelihood-based estimation.

22.7.2 State-Space Representation and Likelihood-Based Estimation

A DSGE model is not automatically an estimable statistical model. In theory, it is a system of equilibrium conditions: households optimize, firms set prices and quantities, policy follows a rule, and shocks evolve over time. To estimate the model, the econometrician must convert those conditions into a probability model for observed data. This requires two steps. First, the rational-expectations system must be solved so that the model implies a stable description of how the economy evolves after shocks. Second, the model's theoretical variables must be linked to measured data series, allowing for measurement error or other differences between the model and the data. Once this link is made, the model can be written in a form that separates unobserved states from observed variables. The Kalman filter can then be used, in the linear-Gaussian case, to evaluate the likelihood and estimate the structural parameters by maximum likelihood or Bayesian methods.

A reliable way to solve linear rational-expectations models was necessary. Blanchard and Kahn (1980) supplied the canonical intuition: predetermined variables (capital, inherited debt) cannot jump today while forward-looking variables (asset prices, inflation expectations, reset prices) can. The model is stable only if jump variables can adjust enough to rule out explosive paths. Too few jump variables prevent stability, while too many leave indeterminacy. The Blanchard–Kahn condition formalizes this – the number of explosive roots must match the number of jump variables, with a rank condition ensuring the jumps can eliminate the explosive directions.

G. Anderson and Moore (1985) and King and Watson (1998) refined the solution theory for linear perfect-foresight and singular systems. Sims (2002) made the solution method more general and computationally convenient by using the QZ decomposition, also known as the generalized Schur decomposition. This numerical method separates the model's expectational difference equations into stable and unstable components. Its advantage is that the researcher does not need to classify every individual variable in advance as predetermined or forward-looking. In many DSGE models, what is predetermined is not a single variable but a linear combination of variables. Sims's *gensys* algorithm handles this general case, checks existence and uniqueness, and, when the model is determinate, returns the stable law of motion used in the state-space representation.

Solution methods also moved beyond first-order linearization. First-order approximations are convenient, but they are certainty-equivalent: they generally suppress the effects of risk on decision rules. Schmitt-Grohé and Uribe (2004) developed second-order perturbation methods, showing that welfare, risk-premium, and precautionary analysis require at least second-order accuracy.

The early likelihood-based estimation literature showed how equilibrium models could be taken directly to the data. Sargent (1989) emphasized that estimation requires a measurement model: different assumptions about how observed data relate to theoretical variables imply different likelihood functions. McGrattan (1994) and Leeper and Sims (1994) carried the likelihood-based approach to RBC models with distortionary taxation and to early policy-usable optimization-based macro models.

Measurement is central to this likelihood program: the mapping between theoretical variables (output, hours, marginal cost) and observed data (national accounts, price indexes, surveys, revised releases) determines which shocks the model infers, which parameters are identified, and how much variation it attributes to theory rather than measurement error.

The linear approximation itself also became part of the empirical specification. Fernández-Villaverde and Rubio-Ramírez (2005) compared likelihood-based estimation using a nonlinear solution with estimation based on a linearized version of the same economy. They showed that nonlinear filtering can deliver substantially different likelihoods and economically meaningful differences in estimated moments, even when the economy appears close to linear. The lesson was that the Kalman filter is appropriate for linear-Gaussian state-space systems, but nonlinear DSGE models generally require simulation-based filters, such as particle filters, if the researcher wants to preserve the nonlinear structure of the model.

By the mid-2010s the workflow had been standardized and codified in handbook form (Fernández-Villaverde, Rubio-Ramírez & Schorfheide, 2016): DSGE models could now be estimated as coherent probability models. The limitation was that full-information likelihood placed heavy weight on every cross-equation restriction, so misspecification and weak identification distorted estimates – the natural response was Bayesian estimation, which uses priors to discipline weakly identified parameters.

22.7.3 The Rise of Bayesian Methods

Likelihood-based DSGE estimation made New Keynesian models genuine statistical models, but medium-scale versions quickly exposed weak identification. Many parameters – preferences, adjustment costs, Calvo probabilities, indexation, policy-rule coefficients, shock variances, and shock persistence– are only weakly identified by limited aggregate data, because different combinations of frictions and shocks can generate similar moments. Bayesian methods responded by combining the likelihood with priors over economically plausible parameter values. The point was not to hide weak identification, but to make explicit the non-sample information used in estimation.

The computational background came from the broader Bayesian econometrics literature. Geweke (1999) emphasized posterior simulation, predictive distributions, and marginal likelihoods as practical tools for models whose posteriors could not be handled analytically. This mattered directly for DSGE estimation. Once a DSGE model is written in state-space form, the Kalman filter can compute the likelihood in the linear-Gaussian case, but the posterior distribution over parameters is usually high-dimensional and nonstandard. Simulation methods, especially Markov chain Monte Carlo, made it possible to report full posterior distributions for parameters, impulse responses, variance decompositions, forecasts, and welfare measures rather than single point estimates.

The early Bayesian DSGE literature showed this could be done in practice. DeJong, Ingram and Whiteman (2000) estimated a stochastic-growth model and made parameter uncertainty central. Schorfheide (2000) introduced loss-function-based model evaluation tied to the policy objects of interest. The major breakthrough was Smets and Wouters (2003). They estimated a medium-scale New Keynesian DSGE model for the euro area using Bayesian methods. The model combined sticky prices, sticky wages, habit formation, investment adjustment costs, variable capacity utilization, monetary policy rules, and multiple structural shocks. The Bayesian framework was essential because the model had too many parameters for the likelihood alone to discipline sharply. Priors anchored economically plausible regions of the parameter space, while the data updated those priors through the likelihood. The result changed the status of DSGE models: they were no longer only calibrated laboratories or small theoretical systems, but estimated forecasting and policy models that could compete with reduced-form benchmarks.

Bayesian methods also created a bridge between structural DSGE models and reduced-form VARs. Del Negro and Schorfheide (2004) introduced the DSGE–VAR, in which a DSGE model supplies a prior for the coefficients of a VAR. The intuition is straightforward. A VAR is flexible but weakly structural; a DSGE is structural but restrictive. The DSGE–VAR centers the VAR around the dynamics implied by the DSGE model, while allowing the data to pull the VAR away from those restrictions. A tight prior pushes the system close to the DSGE; a loose prior allows more reduced-form flexibility. A subsequent DSGE–VAR evaluation of large New Keynesian models (Del Negro, Schorfheide, Smets & Wouters, 2007) concluded that the Smets–Wouters specification was useful for policy but misspecified enough that its restrictions should not be imposed without data-driven relaxation.

By the mid-2000s the workflow had become standardized (An & Schorfheide, 2007; Del Negro & Schorfheide, 2011), and Fernández-Villaverde and Rubio-Ramírez (2007) extended it to nonlinear DSGE likelihoods via the particle filter. By the late 2000s, Bayesian estimation had become the dominant empirical language for medium-scale DSGE models, combining theory, data, and outside evidence in a transparent probability framework that reported uncertainty and allowed formal model comparison. Its weakness was equally clear: priors did real work, marginal likelihoods could be sensitive to prior choices, and posterior precision could create false certainty when all candidate models were misspecified. The tradeoff was now explicit. Structural interpretation required strong restrictions, while empirical fit required enough flexibility to let the data speak.

22.7.4 Medium-Scale DSGE Models: Theory, Estimation, Policy Use

The combination of nominal rigidities, state-space likelihood methods, and Bayesian inference produced the medium-scale DSGE paradigm. These models kept the New Keynesian core – optimizing households and firms, rational expectations, sticky prices, sticky wages, and a monetary-policy rule – but added enough real frictions and structural shocks to confront macroeconomic data. The organizing question of the mid-2000s was whether such models could do two things at once: compete with reduced-form benchmarks in fit and forecasting, while retaining enough structure to support policy analysis.

Christiano et al. (2005), henceforth CEE, supplied the first major template, built to match the impulse responses to identified monetary shocks. The empirical puzzle was that output responds persistently while inflation adjusts only gradually, which a small sticky-price model cannot deliver – strong output growth pushes marginal cost up too quickly. CEE combined sticky prices and wages, habit persistence, investment adjustment costs, variable capital utilization, and a monetary-policy rule. The load-bearing mechanisms were staggered wage contracts and variable utilization, which keep marginal cost from jumping after a shock. This established the medium-scale strategy of matching impulse responses by adding frictions that make propagation realistic.

Smets and Wouters (2007) turned that strategy into the Bayesian workhorse. Their U.S. model explained seven series – output, consumption, investment, hours, real wages, inflation, and the nominal rate, through a New Keynesian DSGE with sticky prices and wages, habit formation, investment adjustment costs, variable capital utilization, and several structural shocks. The model forecast about as well as Bayesian VARs while retaining structural interpretation, and its shock decomposition replaced the simple RBC technology-shock narrative with distinct roles for risk-premium, spending, investment, markup, productivity, and monetary-policy shocks at different horizons.

The same architecture then spread outward into substantive extensions. Galí, López-Salido and Vallés (2007) added rule-of-thumb consumers to explain why private consumption can rise after an increase in government spending, a response that representative-agent Ricardian models struggle to generate. Christiano, Trabandt and Walentin (2017) introduced involuntary unemployment into a medium-scale monetary DSGE by making unemployment a welfare-relevant state rather than treating all nonwork as voluntary leisure. These extensions all responded to the same problem: the representative-agent New Keynesian core was too narrow to explain fiscal transmission, housing cycles, borrowing constraints, and labor-market distress.

DSGE-based policy suites diffused across central banks, including the Federal Reserve Board (Erceg, Guerrieri & Gust, 2006; Edge, Kiley & Laforte, 2008), the European Central Bank (ECB, Christoffel, Coenen & Warne, 2008), and the Riksbank (Adolfson, Laséen, Lindé & Villani, 2007). In these applications, DSGE models became institutional tools that organized forecasts, decomposed historical movements into shocks, and evaluated counterfactual policy paths.

The estimated models also produced new shock-accounting debates. Justiniano, Primiceri and Tambalotti (2010) found that shocks to the marginal efficiency of investment explained a large share of business-cycle variation in output and hours. Justiniano, Primiceri and Tambalotti (2011) then distinguished these marginal-efficiency shocks from investment-specific technology shocks using the relative price of investment goods. The distinction mattered. A fall in the relative price of investment has a clearer technological interpretation; a marginal-efficiency shock is harder to interpret and may capture financial frictions, capital-quality disturbances, or misspecification. The episode illustrated both the power and the fragility of medium-scale DSGE accounting: the model could identify a dominant source of fluctuations, but the economic meaning of that shock depended on measurement and specification.

The main internal critique came from Chari, Kehoe and McGrattan (2009). Many shocks and frictions in estimated New Keynesian models – wage-markup and price-markup shocks in particular, looked like residual wedges needed to fit the data rather than deep disturbances with independent microfoundations. A model can forecast well through flexible shocks, but policy counterfactuals require parameters that remain meaningful when the policy regime changes. The Lucas critique applied back onto medium-scale DSGE models themselves.

Christiano, Trabandt and Walentin (2010); Del Negro and Schorfheide (2013) surveyed the mature policy use of these models for forecasting, historical storytelling, and conditional policy experiments incorporating surveys, nowcasts, and financial spreads. The Great Recession exposed the limits of the pre-crisis benchmark. Many models had weak financial sectors, limited unemployment structure, and were solved around a normal positive-interest-rate steady state, making them poorly suited for a crisis with binding financial constraints and the zero lower bound. But the crisis did not end the medium-scale DSGE program. It redirected it toward financial frictions, occasionally binding constraints, heterogeneous agents, and richer labor-market structure.

By the early 2010s the medium-scale DSGE paradigm had achieved operational maturity but not final credibility. It made structural models estimable, forecastable, and usable in policy institutions, but the shocks and frictions that fit the data often looked less like deep primitives than reduced-form devices. The central question was no longer whether DSGE models could be brought to the data but whether their estimated shocks and frictions were structural enough to support the policy counterfactuals they were being used for.

22.7.5 Stress-Testing DSGE Models: Crisis and Financial Frictions

The medium-scale DSGE paradigm reached operational maturity just before the 2007–2009 financial crisis. Benchmark policy models had rich nominal and real frictions, but they usually placed finance in the background and solved the economy around a normal steady state with a positive nominal interest rate. The crisis made both simplifications costly. Credit spreads, bank balance sheets, collateral values, household debt, liquidity premia, unemployment, and the zero lower bound became central objects rather than peripheral complications. The post-crisis DSGE literature therefore did not begin from scratch. It elevated earlier financial-friction mechanisms and combined them with New Keynesian policy models.

Collateral constraints and financial-accelerator mechanisms became central to the post-crisis DSGE literature. Kiyotaki and Moore (1997) showed how collateral constraints amplify shocks by tightening borrowing limits when asset prices fall, and Bernanke, Gertler and Gilchrist (1999) formalized the financial accelerator as an external-finance premium that rises with falling entrepreneur net worth. Iacoviello (2005) carried the same logic into housing and household borrowing, so that house-price movements shape borrowing capacity and consumption.

The crisis shifted attention from borrowers to intermediaries. Gertler and Kiyotaki (2010) placed leveraged intermediaries under balance-sheet constraints, so that a fall in intermediary net worth contracts credit supply, raises spreads, and depresses activity. Gertler and Karadi (2011) extended this to unconventional policy, in which the central bank intermediates credit directly when private intermediaries cannot, and Cúrdia and Woodford (2010) showed credit spreads enter the optimal Taylor rule when intermediation is costly.

The zero lower bound was a second stress test. Eggertsson and Krugman (2012) showed that debt overhang and forced deleveraging generate a demand-driven slump when nominal rates cannot fall enough: constrained borrowers cut spending, unconstrained savers do not fully offset the fall, and the ZLB prevents monetary policy from restoring demand. Del Negro, Eggertsson, Ferrero and Kiyotaki (2017) extended the framework to emergency liquidity facilities as a concrete unconventional-policy instrument.

Financial observables anchored the empirical side: Gilchrist and Zakrajšek (2012) constructed a corporate credit-spread index and separated expected default risk from an excess bond premium that captures intermediary risk-bearing capacity. Christiano, Motto and Rostagno (2014) then incorporated a financial accelerator into an estimated DSGE and identified 'risk shocks' as a major cyclical driver, though the interpretation of such shocks as primitives remained contested.

The crisis also exposed the limits of linearization. Brunnermeier and Sannikov (2014) provided the leading nonlinear macro-finance benchmark, with financial amplification depending on intermediary balance-sheet state and a 'volatility paradox' through which low measured volatility encourages leverage and later fragility – dynamics that linearized models around a steady state cannot capture.

Estimated DSGE models then tried to explain the Great Recession. Christiano, Eichenbaum and Trabandt (2015) attributed the collapse to financial frictions interacting with the zero lower bound on nominal interest rates, and explained the inflation puzzle through working-capital costs and a productivity fall, and Christiano, Eichenbaum and Trabandt (2016) addressed the labor-market side with a search-and-matching model that derived wage inertia from bargaining rather than imposing Calvo wage stickiness.

The institutional reassessment followed. Lindé, Smets and Wouters (2016) argued that benchmark central-bank DSGE models had difficulty explaining both the depth of the Great Recession and the slow recovery. Adding a zero lower bound, time-varying volatility, and financial-accelerator mechanisms improved the fit, but did not fully solve the policy problem. Central banks now needed models that could handle financial instability, unconventional monetary policy, and macroprudential questions. Christiano, Eichenbaum and Trabandt (2018) offered the internal defense of DSGE modeling, while Stiglitz (2018) gave the external critique. In his view, the crisis revealed deeper problems: representative-agent models, rational-expectations equilibrium, weak financial sectors, and limited attention to information, distribution, and bankruptcy made the pre-crisis paradigm poorly suited to financial crises.

The post-crisis settlement was mixed. What aged well was the DSGE commitment to explicit mechanisms, disciplined counterfactuals, and model-based policy analysis, now studying collateral constraints, credit spreads, intermediary balance sheets, risk shocks, the zero lower bound, and unemployment in a unified equilibrium setting.

What aged less well was the pre-crisis confidence that a representative-agent, log-linear, friction-light model could serve as a general policy laboratory: the shocks and wedges that repaired fit often strained structural interpretation, and crisis dynamics required heterogeneity, balance sheets, nonlinear constraints, and endogenous risk. The standard of adequacy now includes credit booms, leverage, liquidity crises, deep unemployment, and policy at the lower bound.

22.8 Beyond the Representative Agent (2010s–present)

The next step beyond medium-scale DSGE models was not to abandon equilibrium macroeconomics, but to relax its representative-agent and representative-firm shortcuts. In heterogeneous-agent models, the cross-sectional distribution of households, firms, assets, debt, income risk, and production links becomes part of the state of the economy. This changed the transmission mechanism. Monetary policy works not only through a representative intertemporal-substitution margin, but also through labor income, liquidity, debt exposure, asset revaluation, and redistribution; fiscal policy depends on who receives income and how it is financed; and firm or sectoral shocks can become aggregate when large firms, financial constraints, or production networks prevent micro disturbances from washing out. The empirical standard also changed: aggregate moments alone were no longer enough, because the relevant mechanisms had to be disciplined by micro data on earnings risk, consumption responses, wealth portfolios, balance sheets, firm size, and supply-chain linkages. This section follows that shift from household heterogeneity and redistribution channels to firm heterogeneity, production networks, micro-data discipline, and the computational tools that made these models usable.

22.8.1 Household Heterogeneity and Redistribution Channels

Representative-agent DSGE models compress the household sector into a single intertemporal Euler equation. That simplification is useful, but it rules out the distributional mechanisms that became central after the Great Recession: some households are liquidity constrained, some hold most of their wealth in illiquid assets, some face high unemployment risk, and some respond much more strongly to changes in current income than others. The heterogeneous-agent program starts from the opposite premise. Aggregate consumption is not the choice of one representative household; it is the sum of many household decisions made under different asset positions, income risks, and borrowing constraints.

The foundations came from the incomplete-markets literature. Huggett (1993) and Aiyagari (1994) then supplied the canonical incomplete-insurance environment. Households receive uncertain labor income, cannot buy full insurance against bad individual shocks, and can borrow only up to a limit. As a result, they save not only to move consumption across time, but also to protect themselves against future income losses. This precautionary-saving motive changes aggregate saving, the equilibrium interest rate, and the distribution of wealth. Huggett (1993) emphasized the implications for the risk-free rate in an incomplete-insurance economy, while Aiyagari (1994) embedded the same household problem in a production economy where precautionary saving affects the capital stock. Krusell and Smith (1998) added aggregate shocks and found approximate aggregation: in a standard real-business-cycle environment, aggregate dynamics could often be summarized well by aggregate capital rather than by the entire wealth distribution. That result made heterogeneous-agent models tractable, but it also helped delay their use in business-cycle policy analysis. The later heterogeneous-agent New Keynesian (HANK) literature identifies where approximate aggregation breaks down: monetary and fiscal policy depend on who receives income, who holds liquid assets, who is close to a borrowing constraint, and who has a high marginal propensity to consume (MPC).

The first New Keynesian bridge used simple two-type models. Galí et al. (2007) added hand-to-mouth consumers to a sticky-price model of fiscal policy. These households consume their current income instead of fully smoothing consumption against future taxes. This changes the effect of government spending. In a representative-agent Ricardian model, higher government spending tends to lower private consumption because households anticipate future taxes. With hand-to-mouth consumers, government spending can raise labor income and therefore raise private consumption, especially when prices are sticky and deficits postpone the tax burden. These two-type models were deliberately simple, but they isolated the core idea that later HANK models developed fully: policy transmission depends on the distribution of income, liquidity, and access to financial markets.

Werning (2015) clarified that incomplete markets do not automatically overturn representative-agent aggregate-demand logic. Heterogeneity matters when borrowing constraints, liquidity, and income risk generate meaningful differences in MPCs across households.

The full HANK framework of Kaplan, Moll and Violante (2018) embeds an Aiyagari-style household sector inside a New Keynesian model and lets households hold two assets — a liquid asset for spending and an illiquid asset (housing, retirement wealth) that earns more but is costly to adjust. This creates wealthy hand-to-mouth households who may own substantial net worth yet still cut consumption sharply when labor income falls. The transmission of monetary policy changes: instead of operating mainly through intertemporal substitution as in the representative-agent NK model, it works indirectly through labor income, asset returns, inflation, and redistribution across households with different MPCs.

Auclert (2019) turned this insight into a redistribution-accounting framework for monetary policy. A monetary shock affects households through an earnings channel (wages, employment, business income), a Fisher channel (revaluing nominal assets and debts through unexpected inflation), and an interest-rate exposure channel (returns paid and received under different maturity positions). These channels matter when winners and losers have different MPCs, so monetary policy works partly by redistributing income and wealth, not only by moving a representative real rate. The same logic reshaped forward guidance, automatic stabilizers, and fiscal multipliers (McKay, Nakamura & Steinsson, 2016; McKay & Reis, 2016; Auclert, Rognlie & Straub, 2024; Hagedorn, Manovskii & Mitman, 2019).

The computational consolidation came from Auclert, Bardóczy, Rognlie and Straub (2021). Earlier heterogeneous-agent models were hard to solve because the entire household distribution was, in principle, a state variable. Sequence-space Jacobians change the computation. Instead of tracking the whole distribution as a massive state vector, the researcher computes how aggregate objects respond to entire sequences of prices, policies, and shocks around the steady state. A Jacobian measures how aggregate consumption at date t responds to a change in an interest-rate sequence at date s. These derivatives can then be combined to solve for general-equilibrium impulse responses and to estimate models. This made HANK a practical applied framework rather than only a computational frontier.

The broader synthesis (Heathcote, Storesletten & Violante, 2009; Krueger, Mitman & Perri, 2016a; Kaplan & Violante, 2018) is that distribution changes the transmission mechanism: HANK relocates much of monetary transmission from intertemporal substitution to labor income, liquidity, redistribution, and fiscal backing. What aged well was the discipline of modeling household saving and insurance explicitly. What aged less well was approximate aggregation as a justification for treating heterogeneity as a detail behind the aggregate.

22.8.2 Firm Heterogeneity and Production Networks

The firm-side analog of household heterogeneity asks when shocks and distortions at individual firms or sectors show up in aggregate data. The literature asks when input–output links become quantitatively large, especially when firm size is highly unequal, when sectors differ sharply in network centrality, when financial frictions distort capital reallocation, or when markups and other wedges make the economy inefficient.

The early production-network idea goes back to Long and Plosser (1983), which framed sectoral shocks as candidates for aggregate fluctuations through input–output linkages. The firm-dynamics foundation is Hopenhayn (1992), which made entry, exit, productivity dispersion, and the firm-size distribution equilibrium objects.

Gabaix (2011) made firm-level granularity central: when the firm-size distribution is heavily skewed, the largest firms are large enough that their idiosyncratic shocks move aggregate outcomes, so part of what looks like an aggregate productivity shock is the size-weighted sum of shocks to large firms. The network counterpart of Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2012) shows that sectoral shocks become aggregate shocks when the input–output network is sufficiently asymmetric: a shock to a central upstream supplier propagates through many downstream production chains, so micro shocks need not wash out when the economy contains highly central sectors. Natural-experiment evidence, notably Carvalho, Nirei, Saito and Tahbaz-Salehi (2021) on the 2011 Great East Japan Earthquake, then showed propagation both downstream to customers and upstream to suppliers through real production relationships.

Financial frictions add another reason why firm heterogeneity matters: tighter credit conditions disrupt capital reallocation across heterogeneous firms (Khan & Thomas, 2013), and firms with different balance sheets respond differently to monetary shocks (Ottonello & Winberry, 2020).

The modern production-network literature broadened the question from shock propagation to aggregation under distortions. Baqaee and Farhi (2020) generalized efficient-economy aggregation to inefficient economies with markups, taxes, financial frictions, and input–output linkages, and Bigio and La'o (2020); E. Liu (2019) traced how sectoral distortions propagate through production networks and shape industrial policy. These papers shifted the literature from the question 'Can micro shocks move aggregates?' to the broader question 'How do micro shocks and distortions aggregate in an inefficient economy?'

The contribution of this literature is that aggregation itself became a structural problem: aggregate fluctuations depend on the firm-size distribution, the allocation of capital across producers, the input–output network, the location of shocks within that network, and the distortions that shape reallocation. What aged well was the representative-firm benchmark and Hulten's theorem for efficient-economy aggregation with intermediate inputs; what aged less well was the inference that micro shocks and distortions are irrelevant for aggregate dynamics – in economies organized around large firms, financial heterogeneity, and supply chains, the cross section is part of the business-cycle mechanism.

22.8.3 Micro Data as a New Source of Empirical Discipline

Models with heterogeneity required evidence that aggregate time series could not provide. A representative-agent model can be disciplined by aggregate consumption, income, hours, and interest rates. A heterogeneous-agent model also needs to match the distribution of earnings risk, the degree of consumption insurance, the share of households close to liquidity constraints, the distribution of MPCs, the composition of household balance sheets, and the cross-sectional effects of policy. Micro data therefore became a new source of empirical discipline, and the relevant macroeconomic parameters had to be measured in household surveys, administrative records, tax data, transaction data, and linked balance-sheet data. The post-2000 household micro-data literature applied the same discipline to HANK and distributional macroeconomics that the earlier micro-price literature had applied to nominal-rigidity models.

The first discipline came from income-risk and partial-insurance evidence. Guvenen (2009) showed that the estimated persistence of earnings shocks depends on whether workers are allowed to have heterogeneous life-cycle income profiles. If one ignores differences in individual earnings growth, persistent profile heterogeneity can be mistaken for nearly permanent shocks. Blundell, Pistaferri and Preston (2008) linked income risk to consumption by estimating how much income shocks pass through into consumption. Their evidence rejected complete insurance but also showed that households insure transitory shocks more effectively than permanent shocks. Heathcote, Storesletten and Violante (2014) then provided an analytical framework for interpreting joint movements in wages, hours, and consumption as evidence on partial insurance. Administrative earnings records (Guvenen, Karahan, Ozkan & Song, 2021) later showed that earnings shocks are far from Gaussian, displaying negative skewness, fat tails, and asymmetries across age and income groups – features that matter for HANK calibration because precautionary saving depends on tail risk as well as variance.

The second discipline came from direct estimates of marginal propensities to consume. Johnson, Parker and Souleles (2006) used the randomized timing of the 2001 U.S. tax rebates to show that households spent a large share of rebate income quickly, especially households with low income or limited liquid wealth. These results were difficult to reconcile with a frictionless permanent-income model, in which a temporary payment should be spread over a long horizon. Follow-up work found similarly large spending responses to the 2008 stimulus payments (Parker, Souleles, Johnson & McClelland, 2013) and showed that average MPCs hide substantial cross-household heterogeneity tied to mortgage debt and balance-sheet position (Misra & Surico, 2014). Kaplan, Violante and Weidner (2014) gave this fact its portfolio interpretation through the wealthy-hand-to-mouth concept. Many households have positive net worth but little liquid wealth because their assets are tied up in housing or retirement accounts. Fagereng, Holm and Natvik (2021) further confirmed the importance of liquidity using Norwegian administrative data on lottery winnings, showing that liquid wealth, age, and prize size strongly predict how much of a transitory income shock households spend. Together, these papers supplied the empirical targets for two-asset HANK models: the model must match not only the average MPC, but the distribution of MPCs across liquid-wealth, debt, age, and asset-position groups.

The third discipline came from household balance sheets and recession propagation. Saez and Zucman (2016) reconstructed the U.S. wealth distribution from capitalized income tax data and documented the long-run rise in wealth concentration, providing a calibration target for models with realistic wealth distributions. Mian, Rao and Sufi (2013); Mian and Sufi (2018) showed why that distribution matters for business cycles: during the 2006–2009 housing collapse, areas with larger housing-wealth losses experienced larger consumption declines, especially among poorer and more levered households, and the credit-driven household-demand channel generalizes this to credit booms followed by deleveraging-induced demand shortfalls. More broadly, Krueger, Mitman and Perri (2016b) organized the Great Recession evidence around the role of income, wealth, and preference heterogeneity in amplifying aggregate shocks. The lesson was direct: household balance sheets are not background distributional facts; they are state variables for aggregate demand.

The fourth discipline came from micro evidence on monetary transmission. Household-level evidence on monetary transmission (Cloyne, Ferreira & Surico, 2020; Holm, Paul & Tischbirek, 2021; A. L. Andersen, Johannesen, Jørgensen & Peydró, 2023) shows that the consumption response concentrates among mortgage debtors, that liquidity and indirect income effects matter, with sizeable responses also among high-liquidity households, a fact that standard HANK models must explain, and that expansionary policy raises income, wealth, and consumption disproportionately for higher-income households through differences in leverage, risky-asset exposure, and nonlabor income. Monetary policy therefore has distributional incidence as well as aggregate effects.

Micro data changed the empirical standard for macroeconomic models. It was no longer enough for a model to match aggregate consumption, output, inflation, and interest rates. A credible heterogeneous-agent model also had to match earnings-risk dynamics, consumption insurance, wealth concentration, liquidity positions, debt exposure, MPC heterogeneity, and the distributional incidence of policy. Reduced-form micro estimates did not replace structural models, because they usually identify partial-equilibrium or local responses. But they supplied the moments, exposures, and heterogeneity patterns that structural models had to respect. The empirical frontier therefore moved from aggregate fit alone to joint discipline: a model must explain both the macro time series and the cross-sectional behavior of the households and firms that generate them.

22.8.4 Computational Innovations

The heterogeneous-agent turn created a computational problem that representative-agent DSGE models largely avoided. In a representative-agent model, the state of the economy can often be summarized by a few aggregate variables, such as capital, productivity, inflation, and the policy rate. In a heterogeneous-agent model, the distribution of households or firms is itself a state variable. Tomorrow's aggregate consumption, investment, labor supply, and asset prices depend on who holds liquid assets, who is indebted, who is unemployed, which firms are financially constrained, and where shocks hit in the production network. The computational task is therefore not just to solve a larger DSGE model. It is to address an economy in which the cross-section moves over time and feeds back into aggregates.

The classical numerical-methods literature (Maliar & Maliar, 2014; Den Haan, 2010) supplied the pre-deep-learning toolkit – sparse grids, simulation-based projection, integration rules, and global approximation methods – and warned that algorithms producing similar aggregate correlations can imply very different distributional behavior in bad states. In heterogeneous-agent macro, the solution method must track the distribution as accurately as it tracks aggregates.

The first generation of heterogeneous-agent solution methods handled this problem by reducing the dimensionality of the distributional state. Projection–perturbation methods solved the stationary incomplete-markets economy first and then approximated responses to aggregate shocks around that stationary distribution. Parametric-distribution methods summarized the cross-sectional distribution with a finite set of moments or parameters. State-space reduction methods compressed high-dimensional idiosyncratic states into objects that could be handled by perturbation. These methods made heterogeneous-agent models feasible, but they remained technically demanding and often local around a steady state.

The next generation changed the computational frontier. Sequence-space methods transformed the problem in terms of how aggregate variables respond to sequences of prices, policies, and shocks, rather than tracking the full distribution as a massive state vector. Deep-learning methods moved in a different direction: neural networks approximate policy functions, value functions, or equilibrium maps directly. Recent deep-learning approaches show how neural networks can be used to approximate policy functions, value functions, or equilibrium conditions directly (Maliar, Maliar & Winant, 2021; Azinovic, Gaegauf & Scheidegger, 2022; Han, Yang & E, 2022). These tools are especially useful in settings where financial frictions and wealth distributions generate nonlinear, state-dependent dynamics that local representative-agent approximations would miss (Fernández-Villaverde, Hurtado & Nuño, 2023).

What aged well from the heterogeneous-agent program is the core insight that cross-sectional distributions are macroeconomic state variables. The earlier representative-agent benchmark treated differences across households and firms as irrelevant for aggregate dynamics unless they affected aggregate capital or productivity. That approximation is now too strong for many business-cycle questions. Monetary policy depends on household debt, liquidity, and asset exposure; fiscal policy depends on who receives income and who has high marginal propensities to consume; recessions depend on unemployment risk, leverage, and balance-sheet repair; and production shocks depend on firm size, financial position, and network location. The program also aged well empirically. Micro data are now a normal source of discipline for macro models, not an optional supplement. A model that matches aggregate consumption but misses the distribution of liquid wealth, MPCs, leverage, or earnings risk is no longer fully credible.

What aged less well is the early hope that approximate aggregation would make heterogeneity mostly irrelevant for business-cycle analysis. That result remains useful in some neoclassical environments, but it does not generalize to economies with nominal rigidities, incomplete markets, debt, illiquid assets, financial constraints, and policy redistribution. Also aging poorly is the idea that adding heterogeneity automatically solves the identification and misspecification problems of earlier DSGE models. It does not. Heterogeneous-agent models add new state variables, new frictions, new micro moments, and new elasticities. They can fit richer facts, but they can also hide weak identification behind computational complexity. A large HANK or production-network model is not more credible merely because it is more detailed; it is more credible only if its distributional mechanisms are disciplined by micro evidence and robust to alternative specifications.

The lasting lesson is therefore balanced. Heterogeneous-agent models are not a rejection of equilibrium macroeconomics; they are its extension to economies where distribution matters for transmission. Their achievement is to make inequality, liquidity, debt, firm size, and production networks endogenous parts of business-cycle analysis. Their limitation is that richer state spaces do not remove the need for identification, transparency, and empirical discipline. Computation made the heterogeneous-agent program feasible, but computation is not a substitute for economic content. The frontier is now to combine distributional realism, tractable computation, and credible identification in applications where representative-agent models are structurally incapable of answering the question.

22.9 Lessons from the History of Business-Cycle Analysis

22.9.1 What Persisted Across Paradigms

Across the history surveyed in this chapter, the most durable elements of business-cycle analysis were not particular shocks, frictions, or policy doctrines, but recurring standards of explanation. A credible theory of cycles had to say what the impulse was, how it propagated, how the cyclical object was measured, what identified the causal mechanism, and what kind of counterfactual the model could support. The impulse–propagation distinction survived from Slutsky and Frisch into simultaneous-equation models, VARs, RBC models, New Keynesian DSGE models, and heterogeneous-agent frameworks. The language changed from crises, credit, and liquidation to shocks, wedges, frictions, and impulse responses, but the central problem remained the same: explain why disturbances do not simply appear and disappear, but instead spread, persist, and become aggregate.

What aged well, therefore, was the demand for explicit mechanisms rather than mere narratives, for probabilistic reasoning rather than deterministic cycle chronologies, for identification rather than correlation, and for policy counterfactuals disciplined by structure. The NBER measurement tradition aged well as an empirical benchmark; Cowles aged well in making identification central; the Lucas critique aged well in warning against policy counterfactuals based on unstable behavioral equations; the VAR tradition aged well in forcing macroeconomists to take dynamic evidence seriously; and DSGE and HANK models aged well insofar as they kept the ambition of explicit, internally coherent counterfactual analysis.

22.9.2 Recurring Tensions Between Theory and Data

The same history also shows that progress in business-cycle econometrics always came with tradeoffs. Descriptive measurement supplied chronologies and facts, but not causal mechanisms. Cowles-style structural models supplied probability, identification, and policy simulation, but their large-scale versions relied on difficult to defend exclusion restrictions, lag structures, and closure assumptions. RBC models supplied equilibrium discipline and policy-invariant primitives, but asked too much of technology shocks, flexible prices, and a representative-agent. VARs supplied transparent dynamic evidence, but their innovations did not identify economic shocks by themselves. New Keynesian DSGE models restored monetary policy and nominal rigidities to equilibrium macroeconomics, but many estimated shocks and wedges looked more like devices for fitting the data than deep structural primitives.

The lesson is not that one tradition succeeded and the others failed. Rather, each tradition clarified a different failure mode. Too little theory leaves measurement and forecasting without interpretation. Too much structure can produce counterfactuals that are precise but fragile. Too much reliance on aggregate time series can hide heterogeneity, balance sheets, and distributional exposure. Too much computational detail can create the appearance of realism while obscuring identification. The recurring tension is therefore not simply between theory and data, but between empirical fit, structural interpretation, computational tractability, and policy relevance.

22.9.3 Open Questions and Future Directions

The frontier of business-cycle theory is not the search for a final model of the cycle. It is the attempt to build models that can handle the shocks, propagation mechanisms, and data environments that have become central since the financial crisis and the pandemic, while still satisfying the older standards of measurement, identification, and counterfactual discipline.

First, current business-cycle theory must better integrate sectoral supply disturbances, relative-price adjustment, and aggregate inflation. The pandemic cycle and the subsequent inflation episode showed that aggregate demand and aggregate productivity shocks are too coarse for many policy questions. Sectoral demand shifts, supply bottlenecks, energy-price movements, and input–output linkages can jointly determine both output losses and inflation dynamics. This pushes business-cycle theory toward multisector models in which the location of the shock matters as much as its aggregate size.

Second, production networks and international fragmentation are no longer peripheral complications. If shocks propagate through supply chains, trade relationships, energy systems, and large firms, then the representative-firm benchmark is too simple for many modern cycles. The relevant question is not only whether micro shocks can become aggregate, but which nodes, sectors, and countries transmit them, and how policy should respond when stabilization has distributional and sectoral consequences.

Third, the heterogeneous-agent approach remains incomplete. HANK and production-network models have shown that distributions are state variables, but they have not eliminated the older identification problem. The next generation of models must jointly discipline aggregate dynamics and cross-sectional mechanisms: liquid wealth, debt exposure, MPC heterogeneity, earnings risk, firm size, and network position. The new gold standard is a model that matches output, inflation, and the distributional incidence of policy across households or firms.

Fourth, business-cycle models must handle nonlinear transitions, not only small deviations during normal times. Financial crises, liquidity traps, sudden stops, energy shocks, and climate events are not well represented as ordinary Gaussian disturbances around a stable steady state. Occasionally binding constraints, leverage cycles, endogenous risk, and regime shifts require methods that can explain why small shocks sometimes have small effects and sometimes trigger large aggregate responses.

Fifth, the rational-expectations benchmark itself is being reconsidered in heterogeneous and high-dimensional environments. The Lucas critique aged well as a warning that expectations and policy regimes cannot be treated as fixed background conditions. But in models where the state of the economy is a high-dimensional distribution, full model-consistent expectations may impose informational and computational demands that are implausible for actual agents. The open question is how to model expectations in a way that remains disciplined enough for counterfactual analysis without assuming that agents solve the econometrician's full model.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A. & Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, *80*(5), 1977–2016.

Adolfson, M., Laséen, S., Lindé, J. & Villani, M. (2007). Bayesian estimation of an open economy DSGE model with incomplete pass-through. *Journal of International Economics*, *72*(2), 481–511.

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, *109*(3), 659–684.

An, S. & Schorfheide, F. (2007). Bayesian analysis of DSGE models. *Econometric Reviews*, *26*(2–4), 113–172.

Andersen, A. L., Johannesen, N., Jørgensen, M. & Peydró, J.-L. (2023). Monetary policy and inequality. *Journal of Finance*, *78*(5), 2945–2989.

Andersen, L. C. & Carlson, K. M. (1970, April). A monetarist model for economic stabilization. *Federal Reserve Bank of St. Louis Review*, 7–25.

Andersen, L. C. & Jordan, J. L. (1968, November). Monetary and fiscal actions: A test of their relative importance in economic stabilization. *Federal Reserve Bank of St. Louis Review*, *50*(11), 11–24.

Anderson, G. & Moore, G. (1985). A linear algebraic procedure for solving linear perfect foresight models. *Economics Letters*, *17*(3), 247–252.

Anderson, T. W. & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of mathematical statistics*, *20*(1), 46–63.

Aruoba, S. B., Diebold, F. X. & Scotti, C. (2009). Real-time measurement of business conditions. *Journal of Business and Economic Statistics*, *27*(4), 417–427.

Auclert, A. (2019). Monetary policy and the redistribution channel. *American Economic Review*, *109*(6), 2333–2367.

Auclert, A., Bardóczy, B., Rognlie, M. & Straub, L. (2021). Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models. *Econometrica*, *89*(5), 2375–2408.

Auclert, A., Rognlie, M. & Straub, L. (2024). The intertemporal keynesian cross. *Journal of Political Economy*, *132*(12), 4068–4121.

Auerbach, A. J. & Gorodnichenko, Y. (2012). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy*, *4*(2), 1–27.

Azinovic, M., Gaegauf, L. & Scheidegger, S. (2022). Deep equilibrium nets. *International Economic Review*, *63*(4), 1471–1525.

Bańbura, M., Giannone, D., Modugno, M. & Reichlin, L. (2013). Now-casting and the real-time data flow. In *Handbook of economic forecasting* (Vol. 2, pp. 195–237). Elsevier.

Banbura, M., Giannone, D. & Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, *25*(1), 71–92.

Banbura, M., Giannone, D. & Reichlin, L. (2011). Nowcasting with daily data. *European Central Bank, Working Paper*, 18.

Baqaee, D. R. & Farhi, E. (2020). Productivity and misallocation in general equilibrium. *The quarterly journal of economics*, *135*(1), 105–163.

Barro, R. J. (1976). Rational expectations and the role of monetary policy. *Journal of Monetary economics*, *2*(1), 1–32.

Basu, S., Fernald, J. G. & Kimball, M. S. (2006). Are technology improvements contractionary? *American Economic Review*, *96*(5), 1418–1448.

Baxter, M. & King, R. G. (1999). Measuring business cycles: Approximate band-pass filters for economic time series. *Review of Economics and Statistics*, *81*(4), 575–593.

Bernanke, B. S. (1986). Alternative explanations of the money–income correlation. *Carnegie-Rochester Conference Series on Public Policy*, *25*, 49–99.

Bernanke, B. S., Boivin, J. & Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, *120*(1), 387–422.

Bernanke, B. S., Gertler, M. & Gilchrist, S. (1999). The financial accelerator in a quantitative business cycle framework. In J. B. Taylor & M. Woodford (Eds.), *Handbook of macroeconomics* (Vol. 1C, pp. 1341–1393). Amsterdam: Elsevier.

Bernanke, B. S. & Kuttner, K. N. (2005). What explains the stock market's reaction to federal reserve policy? *Journal of Finance*, *60*(3), 1221–1257.

Beveridge, S. & Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics*, *7*(2), 151–174.

Bigio, S. & La'o, J. (2020). Distortions in production networks. *The quarterly journal of economics*, *135*(4), 2187–2253.

Bils, M. & Klenow, P. J. (2004). Some evidence on the importance of sticky prices. *Journal of Political Economy*, *112*(5), 947–985.

Blanchard, O. J. & Kahn, C. M. (1980). The solution of linear difference models under rational expectations. *Econometrica*, *48*(5), 1305–1311.

Blanchard, O. J. & Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, *79*(4), 655–673.

Blanchard, O. J. & Watson, M. W. (1986). Are business cycles all alike? In R. J. Gordon (Ed.), *The american business cycle: Continuity and change*

(Vol. 25, pp. 123–180). Chicago: University of Chicago Press.

Blundell, R., Pistaferri, L. & Preston, I. (2008). Consumption inequality and partial insurance. *American Economic Review*, *98*(5), 1887–1921.

Brunner, K. (1973). *Econometric models of cyclical behavior. volumes i and ii.* JSTOR.

Brunnermeier, M. K. & Sannikov, Y. (2014). A macroeconomic model with a financial sector. *American Economic Review*, *104*(2), 379–421.

Burns, A. F. (1934). *Production trends in the united states since 1870* (No. burn34-1). National Bureau of Economic Research.

Burns, A. F. & Mitchell, W. C. (1946). *Measuring business cycles* (No. burn46-1). National bureau of economic research.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, *12*(3), 383–398.

Canova, F. (1994). Statistical inference in calibrated models. *Journal of Applied Econometrics*, *9*(S1), S123–S144.

Canova, F. (1998). Detrending and business cycle facts. *Journal of Monetary Economics*, *41*(3), 475–512.

Canova, F. & De Nicoló, G. (2002). Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics*, *49*(6), 1131–1159.

Carter, C. K. & Kohn, R. (1994). On gibbs sampling for state space models. *Biometrika*, *81*(3), 541–553.

Carvalho, V. M., Nirei, M., Saito, Y. U. & Tahbaz-Salehi, A. (2021). Supply chain disruptions: Evidence from the great east japan earthquake. *Quarterly Journal of Economics*, *136*(2), 1255–1321.

Chari, V. V., Kehoe, P. J. & McGrattan, E. R. (2008). Are structural VARs with long-run restrictions useful in developing business cycle theory? *Journal of Monetary Economics*, *55*(8), 1337–1352.

Chari, V. V., Kehoe, P. J. & McGrattan, E. R. (2009). New Keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics*, *1*(1), 242–266.

Christiano, L. J. & Eichenbaum, M. (1992). Current real-business-cycle theories and aggregate labor-market fluctuations. *The American Economic Review*, 430–450.

Christiano, L. J., Eichenbaum, M. & Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, *113*(1), 1–45.

Christiano, L. J., Eichenbaum, M. & Trabandt, M. (2016). Unemployment and business cycles. *Econometrica*, *84*(4), 1523–1569.

Christiano, L. J., Eichenbaum, M. & Vigfusson, R. J. (2006). Assessing structural VARs. *NBER Macroeconomics Annual*, *21*, 1–72.

Christiano, L. J., Eichenbaum, M. S. & Trabandt, M. (2015). Understanding the Great Recession. *American Economic Journal: Macroeconomics*, *7*(1), 110–167.

Christiano, L. J., Eichenbaum, M. S. & Trabandt, M. (2018). On DSGE models. *Journal of Economic Perspectives*, *32*(3), 113–140.

Christiano, L. J., Motto, R. & Rostagno, M. (2014). Risk shocks. *American Economic Review*, *104*(1), 27–65.

Christiano, L. J., Trabandt, M. & Walentin, K. (2010). DSGE models for monetary policy analysis. In B. M. Friedman & M. Woodford (Eds.), *Handbook of monetary economics* (Vol. 3A, pp. 285–367). Amsterdam: Elsevier.

Christiano, L. J., Trabandt, M. & Walentin, K. (2017). Involuntary unemployment and the business cycle. *Review of Economic Dynamics*, *24*, 139–221. (Originally circulated as NBER Working Paper 15801, 2010.)

Christoffel, K., Coenen, G. & Warne, A. (2008). *The New area-wide model of the Euro area: A micro-founded open-economy model for forecasting and policy analysis* (ECB Working Paper No. 944). European Central Bank.

Clarida, R., Galí, J. & Gertler, M. (1999). The science of monetary policy: A New Keynesian perspective. *Journal of Economic Literature*, *37*(4), 1661–1707.

Cloyne, J., Ferreira, C. & Surico, P. (2020). Monetary policy when households have debt: New evidence on the transmission mechanism. *Review of Economic Studies*, *87*(1), 102–129.

Cochrane, J. H. (1988). How big is the random walk in GNP? *Journal of Political Economy*, *96*(5), 893–920.

Cogley, T. & Nason, J. M. (1995). Effects of the Hodrick–Prescott filter on trend and difference stationary time series: Implications for business cycle research. *Journal of Economic Dynamics and Control*, *19*(1–2), 253–278.

Cogley, T. & Sargent, T. J. (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S. *Review of Economic Dynamics*, *8*(2), 262–302.

Cooley, T. F. & LeRoy, S. F. (1985). Atheoretical macroeconometrics: A critique. *Journal of Monetary Economics*, *16*(3), 283–308.

Croushore, D. & Stark, T. (2001). A real-time data set for macroeconomists. *Journal of econometrics*, *105*(1), 111–130.

Cúrdia, V. & Woodford, M. (2010). Credit spreads and monetary policy. *Journal of Money, Credit and Banking*, *42*(s1), 3–35.

DeJong, D. N., Ingram, B. F. & Whiteman, C. H. (2000). A Bayesian approach to dynamic macroeconomics. *Journal of Econometrics*, *98*(2), 203–223.

De Leeuw, F. & Gramlich, E. (1968). The federal reserve-mit economic model. *Federal Reserve Bulletin*(Jan), 11–40.

Del Negro, M., Eggertsson, G., Ferrero, A. & Kiyotaki, N. (2017). The Great escape? a quantitative evaluation of the Fed's liquidity facilities. *American Economic Review*, *107*(3), 824–857.

Del Negro, M. & Schorfheide, F. (2004). Priors from general equilibrium models for VARs. *International Economic Review*, *45*(2), 643–673.

Del Negro, M. & Schorfheide, F. (2011). Bayesian macroeconometrics. In J. Geweke, G. Koop & H. van Dijk (Eds.), *Oxford handbook of Bayesian econometrics* (pp. 293–389). Oxford: Oxford University Press.

Del Negro, M. & Schorfheide, F. (2013). DSGE model-based forecasting. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting* (Vol. 2A, pp. 57–140). Amsterdam: Elsevier.

Del Negro, M., Schorfheide, F., Smets, F. & Wouters, R. (2007). On the fit of New Keynesian models. *Journal of Business and Economic Statistics*, *25*(2), 123–143.

Den Haan, W. J. (2010). Comparison of solutions to the incomplete markets model with aggregate uncertainty. *Journal of Economic Dynamics and Control*, *34*(1), 4–27.

Doan, T., Litterman, R. & Sims, C. A. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, *3*(1), 1–100.

Duesenberry, J. S., Eckstein, O. & Fromm, G. (1960). A simulation of the united states economy in recession. *Econometrica: Journal of the Econometric Society*, 749–809.

Duggal, V. G., Klein, L. R. & McCarthy, M. D. (1974). The wharton model mark iii: a modern is-lm construct. *International Economic Review*, 572–594.

Dusenberry, J. S., Fromm, G., Klein, L. R. & Kuh, E. (1965). The brookings quarterly econometric model of the united states. *Chicago: Rand M c Nally Co*.

Edge, R. M., Kiley, M. T. & Laforte, J.-P. (2008). Natural rate measures in an estimated DSGE model of the U.S. economy. *Journal of Economic Dynamics and Control*, *32*(8), 2512–2535.

Eggertsson, G. B. & Krugman, P. (2012). Debt, deleveraging, and the liquidity trap: A Fisher–Minsky–Koo approach. *Quarterly Journal of Economics*, *127*(3), 1469–1513.

Erceg, C. J., Guerrieri, L. & Gust, C. (2006). SIGMA: A new open economy model for policy analysis. *International Journal of Central Banking*, *2*(1), 1–50.

Erceg, C. J., Henderson, D. W. & Levin, A. T. (2000). Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics*, *46*(2), 281–313.

Fagereng, A., Holm, M. B. & Natvik, G. J. (2021). MPC heterogeneity and household balance sheets. *American Economic Journal: Macroeconomics*, *13*(4), 1–54.

Fernández-Villaverde, J., Hurtado, S. & Nuño, G. (2023). Financial frictions and the wealth distribution. *Econometrica*, *91*(3), 869–901.

Fernández-Villaverde, J. & Rubio-Ramírez, J. F. (2005). Estimating dynamic equilibrium economies: Linear versus nonlinear likelihood. *Journal of Applied Econometrics*, *20*(7), 891–910.

Fernández-Villaverde, J. & Rubio-Ramírez, J. F. (2007). Estimating macroeconomic models: A likelihood approach. *Review of Economic Studies*, *74*(4), 1059–1087.

Fernández-Villaverde, J., Rubio-Ramírez, J. F. & Schorfheide, F. (2016). Solution and estimation methods for DSGE models. In J. B. Taylor & H. Uhlig (Eds.), *Handbook of macroeconomics* (Vol. 2A, pp. 527–724). Amsterdam: Elsevier.

Fischer, S. (1977). Long-term contracts, rational expectations, and the optimal money supply rule. *Journal of political economy*, *85*(1), 191–205.

Fisher, I. (1933). The debt-deflation theory of great depressions. *Econometrica: Journal of the Econometric Society*, 337–357.

Friedman, M. (1968). The role of monetary policy. *The American economic review*, *58*(1), 1–17.

Frisch, R. (1933). Propagation problems and impulse problems in dynamic economics." in economic essays in honour of gustav cassel. london: George allan and unwin, ltd.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, *79*(3), 733–772.

Galí, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review*, *89*(1), 249–271.

Galí, J. & Gertler, M. (2007). Macroeconomic modeling for monetary policy evaluation. *Journal of Economic Perspectives*, *21*(4), 25–45.

Galí, J., López-Salido, J. D. & Vallés, J. (2007). Understanding the effects of government spending on consumption. *Journal of the European Economic Association*, *5*(1), 227–270.

Gertler, M. & Karadi, P. (2011). A model of unconventional monetary policy. *Journal of Monetary Economics*, *58*(1), 17–34.

Gertler, M. & Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, *7*(1), 44–76.

Gertler, M. & Kiyotaki, N. (2010). Financial intermediation and credit policy in business cycle analysis. In B. M. Friedman & M. Woodford (Eds.), *Handbook of monetary economics* (Vol. 3A, pp. 547–599). Amsterdam: Elsevier.

Geweke, J. (1999). Using simulation methods for Bayesian econometric models: Inference, development, and communication. *Econometric Reviews*, *18*(1), 1–73.

Ghysels, E., Santa-Clara, P. & Valkanov, R. (2004). *The MIDAS touch: Mixed data sampling regression models* (CIRANO Working Paper No. 2004s-20). CIRANO.

Giannone, D., Lenza, M. & Primiceri, G. E. (2015). Prior selection for vector autoregressions. *Review of Economics and Statistics*, *97*(2), 436–451.

Giannone, D., Reichlin, L. & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, *55*(4), 665–676.

Gilchrist, S. & Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, *102*(4), 1692–1720.

Goodfriend, M. & King, R. G. (1997). The New neoclassical synthesis and the role of monetary policy. In B. S. Bernanke & J. J. Rotemberg (Eds.), *Nber macroeconomics annual* (Vol. 12, pp. 231–283). Cambridge, MA: MIT Press.

Gregory, A. W. & Smith, G. W. (1991). Calibration as testing: inference in simulated macroeconomic models. *Journal of Business & Economic Statistics*, *9*(3), 297–303.

Gürkaynak, R. S., Sack, B. & Swanson, E. T. (2005). Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*, *1*(1), 55–93.

Guvenen, F. (2009). An empirical investigation of labor income processes. *Review of Economic Dynamics*, *12*(1), 58–79.

Guvenen, F., Karahan, F., Ozkan, S. & Song, J. (2021). What do data on millions of

U.S. workers reveal about lifecycle earnings dynamics? *Econometrica*, *89*(5), 2303–2339.

Haavelmo, T. (1944). The probability approach in econometrics. *Econometrica: Journal of the Econometric Society*, iii–115.

Hagedorn, M., Manovskii, I. & Mitman, K. (2019). *The fiscal multiplier* (Tech. Rep.). National Bureau of Economic Research.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*(2), 357–384.

Hamilton, J. D. (2003). What is an oil shock? *Journal of Econometrics*, *113*(2), 363–398.

Hamilton, J. D. (2018). Why you should never use the Hodrick–Prescott filter. *Review of Economics and Statistics*, *100*(5), 831–843.

Han, J., Yang, Y. & E, W. (2022). DeepHAM: A global solution method for heterogeneous agent models with aggregate shocks. *arXiv preprint*. (arXiv:2112.14377)

Hansen, G. D. (1985). Indivisible labor and the business cycle. *Journal of monetary Economics*, *16*(3), 309–327.

Hansen, G. D. & Wright, R. (1992). The labor market in real business cycle theory. *Federal Reserve Bank of Minneapolis Quarterly Review*, *16*(2), 2–12. (Spring)

Hansen, L. P. & Sargent, T. J. (1980). Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and control*, *2*, 7–46.

Harvey, A. C. & Jaeger, A. (1993a). Detrending, stylised facts and the business cycle. *Journal of Applied Econometrics*, *8*(3), 231–247.

Harvey, A. C. & Jaeger, A. (1993b). Detrending, stylized facts and the business cycle. *Journal of applied econometrics*, *8*(3), 231–247.

Hayek, F. A. (1929). *Geldtheorie und konjunkturtheorie* (Vol. 1). Hölder-Pichler-Tempsky ag.

Hayek, F. A. (1931). Prices and production. *George Routledge and Sons*.

Heathcote, J., Storesletten, K. & Violante, G. L. (2009). Quantitative macroeconomics with heterogeneous households. *Annu. Rev. Econ.*, *1*(1), 319–354.

Heathcote, J., Storesletten, K. & Violante, G. L. (2014). Consumption and labor supply with partial insurance: An analytical framework. *American Economic Review*, *104*(7), 2075–2126.

Hodrick, R. J. & Prescott, E. C. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, *29*(1), 1–16.

Holm, M. B., Paul, P. & Tischbirek, A. (2021). The transmission of monetary policy under the microscope. *Journal of Political Economy*, *129*(10), 2861–2904.

Holston, K., Laubach, T. & Williams, J. C. (2017). Measuring the natural rate of interest: International trends and determinants. *Journal of International Economics*, *108*(Supplement 1), S59–S75.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, *60*(5), 1127–1150.

Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*, *17*(5–6), 953–969.

Iacoviello, M. (2005). House prices, borrowing constraints, and monetary policy in the business cycle. *American Economic Review*, *95*(3), 739–764.

Jarociński, M. & Karadi, P. (2020). Deconstructing monetary policy surprises—the role of information shocks. *American Economic Journal: Macroeconomics*, *12*(2), 1–43.

Jevons, W. S. (1879). Sun-spots and commercial crises. *Nature*, *19*(495), 588–590.

Johnson, D. S., Parker, J. A. & Souleles, N. S. (2006). Household expenditure and the income tax rebates of 2001. *American Economic Review*, *96*(5), 1589–1610.

Jordà, O. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, *95*(1), 161–182.

Juglar, C. (1862). *Des crises commerciales et de leur retour périodique en france, en angleterre et aux etats-unis*. Guillaumin et Cie.

Justiniano, A., Primiceri, G. E. & Tambalotti, A. (2010). Investment shocks and business cycles. *Journal of Monetary Economics*, *57*(2), 132–145.

Justiniano, A., Primiceri, G. E. & Tambalotti, A. (2011). Investment shocks and the relative price of investment. *Review of Economic Dynamics*, *14*(1), 102–121.

Kadiyala, K. R. & Karlsson, S. (1997). Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics*, *12*(2), 99–132.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*(1), 35–45.

Kaplan, G., Moll, B. & Violante, G. L. (2018). Monetary policy according to HANK. *American Economic Review*, *108*(3), 697–743.

Kaplan, G. & Violante, G. L. (2018). Microeconomic heterogeneity and macroeconomic shocks. *Journal of Economic Perspectives*, *32*(3), 167–194.

Kaplan, G., Violante, G. L. & Weidner, J. (2014). The wealthy hand-to-mouth. *Brookings Papers on Economic Activity*, *Spring*, 77–153.

Karlsson, S. (2013). Forecasting with Bayesian vector autoregression. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting, vol. 2b* (pp. 791–897). Amsterdam: Elsevier.

Keynes, J. M. (1936). *The general theory of employment, interest and money.* Macmillan.

Khan, A. & Thomas, J. K. (2013). Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy*, *121*(6), 1055–1107.

Kim, C.-J. & Nelson, C. R. (1999). *State-space models with regime switching: Classical and gibbs-sampling approaches with applications, volume 1 of mit press books.* The MIT Press.

Kim, S., Shephard, N. & Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The review of economic studies*, *65*(3), 361–393.

King, R. G., Plosser, C. I. & Rebelo, S. T. (1988a). Production, growth and business cycles: Ii new directions. *Journal of Monetary Economics*, *21*(2-3), 309–341.

King, R. G., Plosser, C. I. & Rebelo, S. T. (1988b). Production, growth and business cycles: I. the basic neoclassical model. *Journal of monetary Economics*, *21*(2-3), 195–232.

King, R. G., Plosser, C. I., Stock, J. H. & Watson, M. W. (1991). Stochastic trends and economic fluctuations. *American Economic Review*, *81*(4), 819–840.

King, R. G. & Rebelo, S. T. (1993). Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, *17*(1–2), 207–231.

King, R. G. & Watson, M. W. (1998). The solution of singular linear difference systems under rational expectations. *International Economic Review*, *39*(4), 1015–1026.

Kitchin, J. (1923). Cycles and trends in economic factors. *The Review of economic statistics*, 10–16.

Kiyotaki, N. & Moore, J. (1997). Credit cycles. *Journal of Political Economy*, *105*(2), 211–248.

Klein, L. R. (1947). The use of econometric models as a guide to economic policy. *Econmetrica: Journal of the Econometric Society*, 111–151.

Klein, L. R. (1950). *Economic fluctuations in the united states, 1921–1941* (No. 11). New York: John Wiley & Sons. (Also published in London by Chapman & Hall)

Klein, L. R. (1964). A postwar quarterly model: description and applications. In *Models of income determination* (pp. 11–57). Princeton University Press.

Klein, L. R. & Goldberger, A. S. (1955). *An econometric model of the united states, 1929–1952* (Vol. 9). Amsterdam: North-Holland Publishing Company.

Klenow, P. J. & Kryvtsov, O. (2008). State-dependent or time-dependent pricing: Does it matter for recent U.S. inflation? *Quarterly Journal of Economics*, *123*(3), 863–904.

Kondratiev, N. (1935). The long-waves in economic in economic life. *Review of Economics and Statistics*, 17.

Koopmans, T. C. (1949). Identification problems in economic model construction. *Econometrica, Journal of the Econometric Society*, 125–144.

Koopmans, T. C. (1950). *Statistical inference in dynamic economic models* (No. 10). Wiley.

Krueger, D., Mitman, K. & Perri, F. (2016a). Macroeconomics and household heterogeneity. In *Handbook of macroeconomics* (Vol. 2, pp. 843–921). Elsevier.

Krueger, D., Mitman, K. & Perri, F. (2016b). Macroeconomics and household heterogeneity. In J. B. Taylor & H. Uhlig (Eds.), *Handbook of macroeconomics* (Vol. 2A, pp. 843–921). North-Holland.

Krusell, P. & Smith, A. A. J. (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy*, *106*(5), 867–896.

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the fed funds futures market. *Journal of Monetary Economics*, *47*(3), 523–544.

Kuznets, S. (1930). Secular movements in production and prices. their nature and their bearing upon cyclical fluctuations (boston: Houghton mifflin co., the riverside press).

Kydland, F. E. & Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of political economy*, *85*(3), 473–491.

Kydland, F. E. & Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, 1345–1370.

Kydland, F. E. & Prescott, E. C. (1990). Business cycles: Real facts and a monetary myth. *Federal Reserve Bank of Minneapolis Quarterly Review*, *14*(2), 3–18.

Kydland, F. E. & Prescott, E. C. (1996). The computational experiment: An econometric tool. *Journal of economic perspectives*, *10*(1), 69–85.

Laubach, T. & Williams, J. C. (2003). Measuring the natural rate of interest. *Review of Economics and Statistics*, *85*(4), 1063–1070.

Leeper, E. M. & Sims, C. A. (1994). Toward a modern macroeconomic model usable for policy analysis. In S. Fischer & J. J. Rotemberg (Eds.), *Nber macroeconomics annual* (Vol. 9, pp. 81–118). Cambridge, MA: MIT Press.

Lindé, J., Smets, F. & Wouters, R. (2016). Challenges for central banks' macro models. In J. B. Taylor & H. Uhlig (Eds.), *Handbook of macroeconomics* (Vol. 2B, pp. 2185–2262). Amsterdam: Elsevier.

Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressions: Five years of experience. *Journal of Business and Economic Statistics*, *4*(1), 25–38.

Liu, E. (2019). Industrial policies in production networks. *The Quarterly journal of economics*, *134*(4), 1883–1948.

Liu, T.-C. (1963). An exploratory quarterly econometric model of effective demand in the postwar us economy. *Econometrica: Journal of the Econometric Society*, 301–348.

Long, J. B. & Plosser, C. I. (1983). Real business cycles. *Journal of political Economy*, *91*(1), 39–69.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of economic theory*, *4*(2), 103–124.

Lucas, R. E. (1973). Some international evidence on output-inflation tradeoffs. *The American economic review*, *63*(3), 326–334.

Lucas, R. E. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, *1*, 19–46.

Lucas, R. E. & Sargent, T. J. (1979). After keynesian macroeconomics. *Federal Reserve Bank of Minneapolis Quarterly Review*, *3*(2), 1–16. (Spring)

Maliar, L. & Maliar, S. (2014). Numerical methods for large-scale dynamic economic models. In K. Schmedders & K. L. Judd (Eds.), *Handbook of computational economics* (Vol. 3, pp. 325–477). North-Holland.

Maliar, L., Maliar, S. & Winant, P. (2021). Deep learning for solving dynamic economic models. *Journal of Monetary Economics*, *122*, 76–101.

Mankiw, N. G. (1989). Real business cycles: A new keynesian perspective. *Journal of economic perspectives*, *3*(3), 79–90.

Mankiw, N. G. (2006). The macroeconomist as scientist and engineer. *Journal of Economic Perspectives*, *20*(4), 29–46.

Mariano, R. S. & Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, *18*(4), 427–443.

McGrattan, E. R. (1994). The macroeconomic effects of distortionary taxation. *Journal of Monetary Economics*, *33*(3), 573–601.

McKay, A., Nakamura, E. & Steinsson, J. (2016). The power of forward guidance revisited. *American Economic Review*, *106*(10), 3133–3158.

McKay, A. & Reis, R. (2016). The role of automatic stabilizers in the us business cycle. *Econometrica*, *84*(1), 141–194.

Mertens, K. & Ravn, M. O. (2013). The dynamic effects of personal and corporate income tax changes in the United States. *American Economic Review*, *103*(4), 1212–1247.

Mertens, K. & Ravn, M. O. (2014). A reconciliation of SVAR and narrative estimates of tax multipliers. *Journal of Monetary Economics*, *68*(Supplement), S1–S19.

Mian, A., Rao, K. & Sufi, A. (2013). Household balance sheets, consumption, and the economic slump. *Quarterly Journal of Economics*, *128*(4), 1687–1726.

Mian, A. & Sufi, A. (2018). Finance and business cycles: The credit-driven household demand channel. *Journal of Economic Perspectives*, *32*(3), 31–58.

Mises, L. V. (1912). *Theorie des geldes und der umlaufsmittel*. Ludwig von Mises Institute.

Misra, K. & Surico, P. (2014). Consumption, income changes, and heterogeneity: Evidence from two fiscal stimulus programs. *American Economic Journal: Macroeconomics*, *6*(4), 84–106.

Mitchell, W. C. (1913). *Business cycles* (Vol. 3). University of California Press.

Mitchell, W. C. (1927). *Business cycles: The problem and its setting* (Vol. 10). National Bureau of Economic Research, Incorporated.

Montiel Olea, J. L., Stock, J. H. & Watson, M. W. (2021). Inference in structural vector autoregressions identified with an external instrument. *Journal of Econometrics*, *225*(1), 74–87.

Moore, G. H. (1961). *Business cycle indicators, volume 1*. National Bureau of Economic Research, Inc.

Morley, J. C., Nelson, C. R. & Zivot, E. (2003). Why are the Beveridge–Nelson and unobserved-components decompositions of GDP so different? *Review of Economics and Statistics*, *85*(2), 235–243.

Murray, C. J. (2003). Cyclical properties of Baxter–King filtered time series. *Review of Economics and Statistics*, *85*(2), 472–476.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: journal of the Econometric Society*, 315–335.

Nakamura, E. & Steinsson, J. (2008). Five facts about prices: A reevaluation of menu cost models. *Quarterly Journal of Economics*, *123*(4), 1415–1464.

Nakamura, E. & Steinsson, J. (2014). Fiscal stimulus in a monetary union: Evidence from US regions. *American Economic Review*, *104*(3), 753–792.

Nakamura, E. & Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: The information effect. *Quarterly Journal of Economics*, *133*(3), 1283–1330.

Nelson, C. R. & Plosser, C. I. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics*, *10*(2), 139–162.

Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review*, *91*(4), 964–985.

Orphanides, A. & van Norden, S. (2002). The unreliability of output-gap estimates in real time. *Review of Economics and Statistics*, *84*(4), 569–583.

Ottonello, P. & Winberry, T. (2020). Financial heterogeneity and the investment channel of monetary policy. *Econometrica*, *88*(6), 2473–2502.

Parker, J. A., Souleles, N. S., Johnson, D. S. & McClelland, R. (2013). Consumer spending and the economic stimulus payments of 2008. *American Economic Review*, *103*(6), 2530–2553.

Phelps, E. S. (1967). Phillips curves, expectations of inflation and optimal unemployment over time. *Economica*, 254–281.

Phelps, E. S. & Taylor, J. B. (1977). Stabilizing powers of monetary policy under rational expectations. *Journal of political Economy*, *85*(1), 163–190.

Phillips, P. C. B. & Jin, S. (2021). Business cycles, trend elimination, and the HP filter. *International Economic Review*, *62*(2), 469–520.

Plagborg-Møller, M. & Wolf, C. K. (2021). Local projections and VARs estimate the same impulse responses. *Econometrica*, *89*(2), 955–980.

Prescott, E. C. (1986). Theory ahead of business-cycle measurement. In *Carnegie-rochester conference series on public policy* (Vol. 25, pp. 11–44).

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, *72*(3), 821–852.

Ramey, V. A. (2011). Identifying government spending shocks: It's all in the timing. *Quarterly Journal of Economics*, *126*(1), 1–50.

Ramey, V. A. & Zubairy, S. (2018). Government spending multipliers in good times and in bad: Evidence from US historical data. *Journal of Political Economy*, *126*(2), 850–901.

Romer, C. D. (1994). Remeasuring business cycles. *The Journal of Economic History*, *54*(3), 573–609.

Romer, C. D. & Romer, D. H. (1989). Does monetary policy matter? a new test in the spirit of friedman and schwartz. In O. J. Blanchard & S. Fischer (Eds.), *Nber macroeconomics annual* (Vol. 4, pp. 121–170). Cambridge, MA: MIT Press.

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

Romer, C. D. & Romer, D. H. (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review*, *100*(3), 763–801.

Rotemberg, J. J. (1982). Sticky prices in the United States. *Journal of Political Economy*, *90*(6), 1187–1211.

Rotemberg, J. J. & Woodford, M. (1997). An optimization-based econometric framework for the evaluation of monetary policy. In B. S. Bernanke & J. J. Rotemberg (Eds.), *Nber macroeconomics annual* (Vol. 12, pp. 297–346). Cambridge, MA: MIT Press.

Saez, E. & Zucman, G. (2016). Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, *131*(2), 519–578.

Sargent, T. J. (1989). Two models of measurements and the investment accelerator. *Journal of Political Economy*, *97*(2), 251–287.

Sargent, T. J. & Sims, C. A. (1977). Business cycle modeling without pretending to have too much a priori economic theory. In C. A. Sims (Ed.), *New methods in business cycle research: Proceedings from a conference* (pp. 45–109). Minneapolis: Federal Reserve Bank of Minneapolis.

Sargent, T. J. & Wallace, N. (1975). " rational" expectations, the optimal monetary instrument, and the optimal money supply rule. *Journal of political economy*, *83*(2), 241–254.

Schmitt-Grohé, S. & Uribe, M. (2004). Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of Economic Dynamics and Control*, *28*(4), 755–775.

Schorfheide, F. (2000). Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics*, *15*(6), 645–670.

Schumpeter, J. A. (1939). Business cycles: A theoretical, historical, and statistical analysis of the capitalist process. *McQraw Hill*.

Sims, C. A. (1972). Money, income, and causality. *American Economic Review*, *62*(4), 540–552.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1–48.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, *36*(5), 975–1000.

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

Slutsky, E. E. (1927). Slozhenie sluchainykh prichin, kak istochnik tsiklicheskikh protsessov. *Voprosy kon"yunktury*, *3*(1), 34–64.

Slutsky, E. E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica: Journal of the Econometric Society*, 105–146.

Smets, F. & Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the Euro area. *Journal of the European Economic Association*, *1*(5), 1123–1175.

Smets, F. & Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review*, *97*(3), 586–606.

Stiglitz, J. E. (2018). Where modern macroeconomics went wrong. *Oxford Review of Economic Policy*, *34*(1–2), 70–106.

Stock, J. H. & Watson, M. W. (1988). Testing for common trends. *Journal of the American Statistical Association*, *83*(404), 1097–1107.

Stock, J. H. & Watson, M. W. (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, *4*, 351–409.

Stock, J. H. & Watson, M. W. (1991). A probability model of the coincident economic indicators. In K. Lahiri & G. H. Moore (Eds.), *Leading economic indicators: New approaches and forecasting records* (pp. 63–89). Cambridge: Cambridge University Press.

Stock, J. H. & Watson, M. W. (1999). Business cycle fluctuations in U.S. macroeconomic time series. In J. B. Taylor & M. Woodford (Eds.), *Handbook of macroeconomics, vol. 1a* (pp. 3–64). Amsterdam: Elsevier.

Stock, J. H. & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, *20*(2), 147–162.

Stock, J. H. & Watson, M. W. (2006). Forecasting with many predictors. In G. Elliott, C. W. J. Granger & A. Timmermann (Eds.), *Handbook of economic forecasting, vol. 1* (pp. 515–554). Amsterdam: Elsevier.

Stock, J. H. & Watson, M. W. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In J. B. Taylor & H. Uhlig (Eds.), *Handbook of macroeconomics, vol. 2a* (pp. 415–525). Amsterdam: Elsevier.

Stock, J. H. & Watson, M. W. (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *Economic Journal*, *128*(610), 917–948. (Sargan Lecture, Royal Economic Society)

Summers, L. H. (1986). *Some skeptical observations on real business cycle theory*. Harvard Institute of Economic Research.

Taylor, J. B. (1980). Aggregate dynamics and staggered contracts. *Journal of Political Economy*, *88*(1), 1–23.

Tenreyro, S. & Thwaites, G. (2016). Pushing on a string: Us monetary policy is less powerful in recessions. *American Economic Journal: Macroeconomics*, *8*(4), 43–74.

Theil, H. (1953). *Estimation and simultaneous correlation in complete equation systems* (Tech. Rep.). Centraal Planbureau.

Tinbergen, J. (1939a). Statistical testing of business cycle theories: Part i: A method and its application to investment activity.

Tinbergen, J. (1939b). Statistical testing of business cycle theories: Part ii: Business cycles in the united states of america, 1919-1932.

Tinbergen, J. (1940). Econometric business cycle research. *The Review of Economic Studies*, *7*(2), 73–90.

Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, *52*(2), 381–419.

Watson, M. W. (1986). Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, *18*(1), 49–75.

Werning, I. (2015). *Incomplete markets and aggregate demand* (Tech. Rep.). National Bureau of Economic Research.

Wicksell, K. (1898). *Geldzins und güterpreise: eine studie über die den tauschwert des geldes bestimmenden ursachen*. Fischer.

Wu, J. C. & Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, *48*(2–3), 253–291.

Yule, G. U. (1927). On a method of investigating periodicities disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *226*(636-646), 267–298.

Yun, T. (1996). Nominal price rigidity, money supply endogeneity, and business cycles. *Journal of Monetary Economics*, *37*(2), 345–370.

Zellner, A. & Theil, H. (1962). Three-stage least squares: Simultaneous estimation of simultaneous equations. *Econometrica*, *30*(1), 54–78.