Theory of Income 2

Paper's Summary fo the Final

1 **Aiyagari, S. Rao. 1994. Uninsured idiosyncratic risk and aggregate saving.

We provide an exposition of models whose aggregate behavior is the result of market interaction among a large number of agents subject to idiosyncratic shocks. This class of models involves a considerable amount of individual dynamics, uncertainty, and asset trading which is the main mechanism by which individuals attempt to smooth consumption. However, aggregate variables are unchanging. This contrasts with representative agent models in which individual dynamics and uncertainty coincide with aggregate dynamics and uncertainty.

We use a standard growth model modified to include a role for uninsured idiosyncratic risk and liquidity/borrowing constraints. This is done by having a large number of agents who receive idiosyncratic labor endowment shocks that are uninsured.

As a result of this market incompleteness in combination with the possibility of being borrowing constrained in future periods, agents accumulate excess capital in order to smooth consumption in the face of uncertain individual labor incomes. The interest rate, aggregate capital, and the wealth distribution are all jointly determined in the presence of precautionary motives and borrowing constraints.

We use such a model to study the quantitative importance of individual risk for aggregate saving. This study is motivated by the debate concerning the sources of aggregate capital accumulation, in particular, the suggestion that precautionary saving may be a quantitatively important component of aggregate saving.

In a steady state consumers face a constant interest rate because the shocks are purely idiosyncratic and, hence, there is no aggregate uncertainty. For a given interest rate optimal individual saving behavior leads to a distribution of agents with different levels of assets reflecting different histories of labor endowment shocks. Aggregation implies some level of per capita assets. In a steady-state equilibrium

the per capita amount of capital must equal the per capita asset holdings of consumers, and the interest rate must equal the net marginal product of capital. These features in combination explain why the interest rate is necessarily less than the time preference rate and, hence, the aggregate capital stock and the saving rate are necessarily greater than under certainty (equivalently, complete markets).

The quantitative results of this paper suggest that the contribution of uninsured idiosyncratic risk to aggregate saving is quite modest.

In contrast to representative agent models, it turns out that access to asset markets is quite important in enabling consumers to smooth out earnings fluctuations

The model is also consistent, at least qualitatively, with certain features of income and wealth distributions. The distributions are positively skewed (median $<$ mean), the wealth distribution is much more dispersed than the income distribution, and inequality as measured by the Gini coefficient is significantly higher for wealth than for income.

This class of models can also differ from the infinite-lived agent complete markets model on some important policy issues. For instance, with complete markets, dynamic optimal factor taxation leads to the result that the capital income tax should be zero in the long run. However, with idiosyncratic shocks and incomplete markets the capital income tax is strictly positive even in the long run.

2 *Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song. 2015. What do data on millions of U.S. workers reveal about life-cycle earnings risk ?

An enduring question for economists is whether these wide-ranging labor market histories, experienced by a diverse set of individuals, display sufficiently simple regularities that would allow researchers to characterize some general properties of earnings dynamics over the life cycle.

How well is the probability distribution of earnings shock approximated by a lognormal distribution, an assumption often made out of convenience ?

The goal of this paper is to characterize the most salient aspects of life-cycle earnings dynamics using a large and confidential panel data set from the U.S. Social Security Administration.

This analysis reaches two broad conclusions. First, the distribution of individual earnings shocks displays important deviations from lognormality. Second, the magnitude of these deviations varies greatly both over the life cycle and with the earnings level of individuals.

First, starting with the first moment, we find that average earnings growth over the life cycle varies strongly with the level of lifetime earnings : the median individual by lifetime earnings experiences an earnings growth of 38 percent from ages 25 to 55, whereas for individuals in the 95th percentile, this figure is 230 percent.

Second, turning to the third moment, we see that earnings shocks are negatively skewed, and this skewness becomes more severe as individuals get older or their earnings increase (or both). Furthermore, this increasing negativity is due entirely to upside earnings moves becoming smaller from ages 25 to 45, and to increasing "disaster" risk (the risk of a sharp fall in earnings) after age 45. Although these implications may appear quite plausible, they are not captured by a lognormal specification, which implies zero skewness.

Third, studying the fourth (standardized) moment, we find that earnings changes display very high kurtosis. There are far more people with very small earnings changes in the data compared with what would be predicted by a normal density. Furthermore, this average kurtosis masks significant heterogeneity across individuals by age and earnings.

Fourth, we characterize the dynamics of earnings shocks by estimating nonparametric impulse response functions conditional on the recent earnings of individuals and on the size of the shock that hits them. We find two types of asymmetries. One, fixing the shock size, positive shocks to high-earnings individuals are quite transitory, whereas negative shocks are very persistent ; the opposite is true for low-earnings individuals. Two, fixing the earnings level of individuals, the strength of mean reversion differs by the size of the shock : large shocks tend to be much more transitory than small shocks.

A corollary to these findings is that the workhorse model in the literature (a persistent AR(1), or random walk, process plus a transitory shock with normal innovations) fails to match most of the prominent features of the earnings data documented in this paper.

Our analysis of the life-cycle earnings histories of millions of U.S. workers has reached two broad conclusions. First, the higher-order moments of individual earnings shocks display clear and important deviations from lognormality. In particular, earnings shocks display strong negative skewness (what can be viewed as

individual disaster shocks) and extremely high kurtosis The high kurtosis implies that in a given year, most individuals experience very small earnings shocks, few experience middling shocks, and a small but non-negligible number experience extremely large shocks.

The second conclusion of this analysis is that these statistical properties of earnings shocks change substantially both over the life cycle and with the earnings level of individuals. For example, our estimates of the stochastic process show that low-income individuals experience very large earnings shocks with low persistence, whereas high-income individuals experience shocks that are very persistent but with much lower volatility.

We have also estimated impulse response functions of earnings shocks and found significant asymmetries : positive shocks to high-earnings individuals are quite transitory, whereas negative shocks are very persistent ; the opposite is true for low-earnings individuals. While these statistical properties are typically ignored in quantitative analyses of life-cycle models, they are fully consistent with search-theoretic models of careers over the life cycle.

A broader message of this paper is a call to reconsider the way researchers approach the study of earnings dynamics. The covariance matrix approach that dominates current work is too opaque and a bit mysterious : it is difficult to judge what it means to match or miss certain covariances in terms of their economic implications. With the current trend toward the increasing availability of very large panel data sets, we believe that researchers' priority in choosing methods should shift from efficiency concerns to transparency. The approach adopted here is an example of the latter, and we believe that economists can better judge what each moment implies for economic questions.

3 *Deaton, Angus. 1991. Saving and liquidity constraints.

This paper is concerned with the optimal intertemporal consumption behavior of consumers who are restricted in their ability to borrow to finance consumption. The restriction is not a symmetric one. Nothing prevents these consumers from saving and accumulating assets, and under some circumstances they will find it desirable to do so.

Limited borrowing opportunities may help to explain the observed patterns of household wealth holdings as well as the fact that consumption appears to track household income quite closely over the life-cycle. Most versions of life-cycle models predict a dissociation of consumption from income, and the existence of substantial

asset accumulations at least at some points in the life cycle. The validity of these predictions has been challenged. In particular, it is clear that most households in the U.S. hold very few assets.

In this paper I consider the behavior of relatively impatient consumers, who prefer consumption now to consumption later. I assume that consumers are "prudent" and have a precautionary demand for saving. Precautionary motives interact with liquidity constraints because the inability to borrow when times are bad provides an additional motive for accumulating assets when times are good, even for impatient consumers.

My general procedure is to start from a simple stochastic process for labor income, and to derive, from that process, the appropriate policy rule for consumption given that borrowing is not allowed, or at least cannot exceed some fixed limit. I shall discuss whether it is possible to build a representative agent model of a liquidity constrained consumer that could account for the main features of the aggregate time-series data in the U.S. But my more fundamental concern is to characterize the type of microeconomic behavior that borrowing constraints might produce. The approach is one of partial equilibrium.

The analysis shows that, in the presence of borrowing restrictions, the behavior of saving and asset accumulation is quite sensitive to what consumers believe about the stochastic process generating their incomes. In the simplest case, when incomes are stationary, independently and identically distributed over time, the consumer saves and dissaves in order to smooth consumption in the face of income uncertainty.

Positive serial correlation in the income process diminishes both the desirability and the feasibility of using assets in this way. In the limit, when income is a random walk, with or without drift, it turns out that those who wish to borrow but cannot do so typically can do no better than consume their incomes.

I also investigate the consequences of borrowing restrictions in an environment in which income growth is stationary, but where the growth rates mimic aggregate data and are positively serially correlated. These models produce what may appear to be the paradoxical result that, when consumers follow the optimal consumption policy, savings is contra-cyclical.

In reality, microeconomic income processes are very different from their macroeconomic aggregates, so that while individual consumers share in the general growth, the variance in their incomes is dominated by idiosyncratic components, some permanent, some transitory.

I construct a simple model in which individual income growth is negatively autocorrelated, aggregate income growth is positively autocorrelated, and aggregate saving is procyclical.

4 **Barro, Robert J. 1979. On the determination of the public debt.

The "Ricardian' equivalence theorem on public debt' says that shifts between debt and tax finance for a given amount of public expenditure would have no first-order effect on the real interest rate, volume of private investment, etc.

Proponents of the Ricardian view that the choice between debt and taxes does not matter are left with an embarrassing absence of a theory of public debt creation. This paper develops a simple theory of "optimal' public finance that identifies some factors that would influence the choice between taxes and debt issue. The model accepts the Ricardian invariance theorem as a valid first-order proposition but introduces some second-order considerations involving the "excess burden" of taxation to obtain a determinate (optimal) amount of debt creation.

The theoretical model is used to formulate several testable propositions concerning the determination of public debt issue. Principal hypotheses involve the positive effect on debt issue of temporary increases in government spending (especially important during war and postwar periods), the negative effect of temporary increases in income-that is, a countercyclical response of debt issue, and a one-to-one effect of the expected inflation rate on the growth rate of nominal debt. The theory also implies that the growth rate of debt would be independent of the debt-income ratio and would be affected at most in a minor way by the level of government expenditure.

The hypotheses are tested using the time-series data on public debt issue in the United States since World War I. The results are basically in accord with the underlying theory.

5 **Lucas, Robert E., Jr. and Nancy L. Stokey. 1983. Optimal fiscal and monetary policy in an economy without capital.

This paper is an application of the theory of optimal taxation to the study of aggregative fiscal and monetary policy.

Qne 'reason' for the time-inconsistency of optimal policies is the classical issue of the 'capital levy'. In the Ramsey framework, with lump-sum taxes assumed unavailable, it is best to focus excise taxes on goods that are inelastically supplied or demanded, to tax 'pure rents'. In a dynamic setting, goods produced in the past, capital, always have this quality and the returns to such goods are thus 'optimally taxed away.

In the present paper, we consider only economies without capital of any form. Private and government consumption goods are assumed to be produced under constant returns to scale using labor as the only input, and government consumption is taken to follow an exogenously given stochastic process. Moreover, the analysis is conducted in a neoclassical framework, thus precluding any countercyclical role for fiscal or monetary policy.

We consider a barter economy and assume that in each period the current government has full control over current tax rates, the issue of new debt, and the refinancing (at market prices) of old debt. However, ir takes as fully binding the debt commitments made by its predecessors. Our main finding is that with debt commitments of a sufficiently rich maturity structure, an optimal policy, if one exists, can be made time consistent.

This paper has been concerned with the structure and time-consistency of optimal tax policy in two multiperiod economies : a pure barter system and a monetary economy, both without capital goods. In each case, the government had to choose a method of financing an exogenous stochastic equence of government expenditures. Current consumption goods and a complete set of contingent claim securities were assumed to be traded in each period.

We showed that the optimal tax policy is time-consistent, provided that fully binding debt of a suficiently rich maturity and risk structure can be issued, and that the optimal debt policy is unique.

Money, in the form of currency, was introduced via a transactions demand, along with nominally-denominated debt. The analogy between the monetary economy and a two-good barter system permitted us to apply the previous analysis. The analogy with the barter system broke down when time-consistency was considered.

Time-consistency can be achieved only if monetary policy is pre-set to maintain a specified path of nominal prices. somewhat surprisingly, this same effect cannot be achieved through a pre-set path for the quantity of money, since the interaction of fiscal and monetary policy permits tax policies to alter the effects on prices of any given monetary policy.

Some form of institutional commitment is essential for the implementation of fiscal and monetary policies that have desirable effects under the usual welfare-economic criteria.

6 **Chamley, Christophe. 1986. Optimal taxation of capital income in general equilibrium with infinite lives.

There is considerable debate in the tax policy community over whether or not to tax the income from capital and at what rates.

The studies of second best taxation indicate that the optimal capital income tax rates are not zero, except under quite restrictive assumptions about individual preferences. The main result of this paper is that when individuals have infinite lives and a utility function of a fairly general form, the optimal tax rate on capital income does tend to zero in the long run.

Previous studies of second best taxation in dynamic general equilibrium found different results because they relied on life-cycle models with limited scope for intertemporal substitution. In these earlier models, the individual's horizon was finite.

I first consider the case where consumption levels have no impact on preferences in future periods. I establish the zero taxation of capital income in the steady state for this class of utility functions : in the intertemporal program of second best taxation, the tax rate on capital income tends to zero as the economy tends to a steady state, or a balanced growth path.

This result is first established in a heuristic model of intertemporal general equilibrium. A unique good is produced with capital and labor ; it can be consumed or used to increase the capital stock. There is no lump-sum taxation ; taxes on income from capital and labor are linear ; and the tax rates can be optimized in each period. Finally, all agents are identical. These assumptions are not essential for the derivation of the result. The main property of the model which is used in the proof is the equality between the private and the social discount rate in the long run.

The utility function is extended to allow for an effect of present decisions on future preferences : The utility function is separable between sequences of consumption programs in k consecutive periods, where k is an arbitrary fixed number. In this analysis, capital taxation is also equal to zero in the steady state, given

economies with many consumption goods and with many agents of heterogeneous endowments and preferences.

An essential part of the argument in this paper is the existence of a stable steady state for the dynamic path of second best. Therefore, Section 3 analyzes this problem for additively separable utility functions. Two results emerge.

First, the steady state is locally stable when one of two independent, sufficient conditions is satisfied. The first of these is a restriction on the shape of the utility function. The second is that the efficiency cost of taxation is relatively small. The latter condition, seems to be applicable for a wider class of models. It implies that the dynamic path of second best is close to the path of first best. When the steady state is stable in the first best, it is also stable in the second best by a continuity argument.

Second, I find that the second best level of the capital income tax is fully described along the economy's dynamic transition path. At first, it is as large as constraints permit. It then abruptly switches, in finite time, to zero. This property is specific to the utility function which is considered here.

7 **Chari, V.V., Juan Pablo Nicolini, and Pedro Teles. 2016. More on the optimal taxation of capital.

In this paper we take the same Ramsey approach to the optimal taxation of capital, in that the tax system is exogenously given, but enlarge the set of instruments to include other taxes widely used in practice in developed economies such as dividend, consumption and wealth taxes. We refer to a tax system with this enlarged set of tax instruments as a rich tax system.

We begin by studying the standard neoclassical growth model with a representative agent. We show that with a rich tax system, capital should not be taxed in the steady state. Along the transition capital may be taxed or subsidized.

We then consider a class of economies that are standard in the macroeconomics literature and show that with these preferences, future capital should never be taxed, except possibly for a one period transition. We then consider heterogeneous agent economies in which agents differ in their initial wealth. We show that the representative agent results also hold in those economies with heterogeneous agents.

Our results differ from those that presume that capital taxes should be high for some lenght of time because we allow for a rich tax system, while that literature considers a restricted tax system.

We also analyze restricted tax systems like those in the literature described earlier. With taxes on capital and labor only and with capital taxes restricted to be less than 100%, we recover the results in Chamley and Straub and Werning that capital should be taxed for some, possibly innite, length of time.

The conventional view of Ramsey equilibrium in dynamic economies is that the government chooses policies in period zero and commits to these policies thereafter. Our results hold under this conventional view. We go on to show that our results hold under an alternative view. In this alternative view, the government in each period has partial commitment in the sense that it can commit to one period returns on assets in utility terms. That is, the government in the following period is free to choose policies as it desires but must respect the previously committed return constraints.

The idea behind this notion of partial commitment begins by noting that returns on assets must satisfy intertemporal Euler equations on the equilibrium path. Requiring governments in any period to satisfy these restrictions allows governments to choose policies as they see t but ensures that both on an off the equilibrium path the policies do not induce agents to regret their past choices. In particular, governments can deviate from equilibrium choices, but these deviations must have the property that agents do not wish that they could go back to the previous period and change the previous choices.

We show that the Markov equilibrium in this set up coincides with the commitment equilibrium with restrictions on the value of wealth.

We go on to study partial commitment to future taxes, rather than returns on assets. We argue that Markov equilibria do not coincide with the commitment equilibrium. The reason is that commitment outcomes face a time inconsistency problem. In particular with single period debt, the government has a strong incentive to choose policies so as to reduce the value of inherited debt.

We briey analyze an economy with heterogeneous agents and show that our representative agent results hold in such economies. An interesting feature of heterogeneous agent economies is that even if initial policies are unrestricted, the Ramsey equilibrium could distort intratemporal margins in order to achieve redistributive roles.

Once we abstract from the initial confiscation of capital, what matters are consumption and labor elasticities. In the steady state they are constant, so capital

taxes should be zero in the steady state. For standard macro preferences, those elasticities are always constant, so that (future) capital should never be taxed. These results hold with heterogeneous agents.

Future capital should not be taxed for standard preferences, because taxing it imposes different taxes on different consumption goods and on labor in different periods, when instead uniform consumption and labor taxation is optimal.

8 **Freeman, Scott, and Finn E. Kydland. 2000. Monetary aggregates and output.

In this paper we ask whether the endogenous nature of monetary aggregates may account in a quantitatively plausible way for the observed procyclical movement of the nominal money stock.

We adapt an endogenous money-multiplier model into an otherwise standard model of a business cycle set off by real disturbances. In deliberate contrast to monetary models that create a money-output link using sticky prices or fixed money holdings, all prices and quantities are assumed to be fully flexible.

The model will be calibrated to meet long-run features of the U.S. economy (including monetary features) and then subjected to shocks to the technology level following a random process like that observed in U.S. data. The model's predicted business cycle frequency correlations, of both real and nominal variables, are then compared to those of the U.S. data.

We find that the model's predicted business-cycle frequency correlations share the following features with U.S. data :

— MI is positively correlated with real output.
— The money multiplier and deposit-to-currency ratio are positively correlated with real output.
— The price level is negatively correlated with output.
— The correla tion of MI with contemporaneous prices is sub stantially weaker than the correlation of MI with real output.
— Correlations among real variables are essentially unchanged under different monetary-policy regimes.
— Real money balances are smoother than money demand equations would predict.

We assume that consumption goods can be purchased using either currency or bank deposits. Two transaction costs are assumed. One is a cost of acquiring money

balances, necessary to determine the demand for money and make endogenous the velocity of money. The other is a fixed cost of using deposits, necessary to determine the division of money balances into currency and interest-bearing deposits. Faced with these two costs and other factors that may vary over the business cycle, households make decisions that determine the velocity of money and the money multiplier.

The key feature of the model is the endogeneity of the money supply that results from the households' choices of the composition of their money balances in response to variables that fluctuate over the business cycle. These endogenous monetary responses yield not only the sought-after money-output correlation but also the sticky prices and money balances that other models of money and output impose by assumption.

We do not offer this analysis as a definitive affirmation of complete monetary neutrality. In deed, in our model economy, shocks to required reserves or serially correlated shocks to the monetary base have an influence on output.

9 *Lucas, Robert E. and Juan Pablo Nicolini. 2015. On the stability of money demand.

The failure of Lehman Brothers in September, 2008 immediately led to a severe banking panic, a rush by banks to exchange privately issued cash substitutes for government issued or government guaranteed cash. The Federal Reserve responded to this situation by increasing the level of bank reserves from some $40 billion on September 1 to $800 billion by New Years Day. This single action was surely the main factor in the resolution of the liquidity crisis by early 2009 and the ending of the decline in production after two quarters.

It is a remarkable feature of these events that none of the leading macroeconometric models, including the model in use by the Fed itself, had anything to contribute to the analysis of this liquidity crisis or of the Fed's response to it. None of these models had any role for bank reserves or for any other monetary aggregate or measure of liquidity. Central bankers, as always, used short interest rates as the only indicator of the stance of monetary policy but sometime in the 1990s they were joined by most influential monetary economists. A broad consensus was reached that no measure of "liquidity" in an economy was of any value in conducting monetary policy.

There were understandable reasons behind this consensus. Long standing empirical relations connecting monetary aggregates like M1, M2 and the monetary

base to movements in prices and interest rates began to deteriorate in the 1980s and have not been restored since. One objective in this paper is to offer a diagnosis of this empirical breakdown. A second is to propose a fix, to construct a new monetary aggregate that offers a unified treatment of monetary facts preceding and following 1980.

To do this we need to get behind such broad aggregates as M1 and M2 and model the role of currency and different kinds of deposits in the payment system in an explicit way. We consider the distinct roles of currency, reserves, and commercial bank deposits. The model proposes a banking "technology" that rationalizes the adding-up of different assets to form aggregates like M1. It treats currency and demand deposits as distinct assets and can readily be adapted to include other forms of liquidity.

The model rationalizes our use of a new, single monetary aggregate -we call it NewM1- that coincides with M1 prior to 1982 and includes the newly created MMDAs for the years since.

We calibrate the model and compare its implications to U.S data for the free interest rates period (1915-1935 and 1983-2012). We show the model matches the behavior of the aggregate NewM1 remarkably well for the entire period.

The ratio of M1 to GDP for the United States from 1915 to 1980 against the 3 month Treasury Bill rate had a clear negative relation, which has been documented many times in empirical studies of money demand. These studies have typically ignored the distinction between the currency and deposit components of M1. In a sense, there was no need to address the composition issue since currency remained close to 25% of deposits during the period.

The relationship between the ratio of M1 to GDP for the United States against the 3 month Treasury Bill broke down in the 1980s. The breakdown in M1 is associated with a breakdown in the behavior of deposits, not in currency.

The early eighties were a hectic period in terms of regulatory changes. Regulation Q was first relaxed slightly in 1980, when banks were allowed to pay limited interest on personal checking accounts (NOW accounts). In 1982 banks were allowed to issue interest-paying money market deposit accounts (MMDA), which could be held by some corporations as well as by households.

The Fed included NOW accounts in M1, together with the traditional zero-interest checking accounts. However, it included MMDA with other savings accounts, which are part of M2.

How did all these regulatory changes affect household decisions regarding the relative desirability of the different components of M1 ? Our candidate theory will be that the changes in the regulation in the early eighties substantially increased the availability of close substitutes for deposits, but not for cash. Once these close substitutes are allowed, our theory implies that they should be included in the monetary aggregate together with cash and traditional checking accounts.

10 *Krishnamurthy, Arvind, and Annette Vissing-Jorgensen 2012. The aggregate demand for treasury debt.

Money, such as currency or checking accounts, offers a low rate of return relative to other assets. The reason is that money is (1) a medium of exchange for buying goods and services, (2) of high liquidity, and (3) of extremely high safety in the sense of offering absolute security of nominal repayment. Investors value these attributes of money and drive down the yield on money relative to other assets.

We argue that Treasury bonds have some of the same features as money, namely (2) and (3), and that this drives down the yield on Treasuries relative to assets that do not to the same extent share these features.

When the supply of Treasuries is low, the value that investors assign to the liquidity and safety attributes offered by Treasuries is high. As a result the yield on Treasuries is low relative to the yield on the Aaa corporate bonds which offer less liquidity and safety. The opposite applies when the supply of Treasuries is high.

We examine the yield spread between a pair of assets which are different only in terms of their liquidity, as well as the yield spread between a pair of assets which are different only in terms of their safety. Under the hypothesis that liquidity/safety are priced attributes, the yield spread between these pairs of assets should reflect the equilibrium price of liquidity/safety. We show that changes in Treasury supply affect each of these yield spreads. The results indicate that Treasuries offer liquidity and safety so that changes in the supply of Treasuries separately change the equilibrium price of liquidity and safety.

The second type of evidence that Treasuries share (2) and (3) with money relates the quantity of money to the quantity of Treasuries. We show that when the supply of Treasuries falls, the supply of bank-issued money (specifically, M2

minus M1) which offers (2) and (3), rises. We show that the channel underlying this response is that a reduction in the supply of Treasuries increases the prices of liquidity and safety, lowering the yield on bank deposits, and inducing the banking sector to issue more deposits.

The government collects seignorage from the liquidity and safety attributes of Treasuries.

The demand for money stems from demand for liquidity, safety, and a medium-of-exchange. It is likely that there is independent variation in the demand and supply of each of these attributes. For example, during a financial crisis, the demand for liquidity and safety in particular may rise. On the other hand, the extant literature on money demand assumes that money reflects the price of a single attribute. We discuss how to use the yields on Treasuries as well as other safe/liquid assets to recover the underlying demands for the different attributes of money.

Our findings imply that Treasury interest rates are not an appropriate benchmark for "riskless" rates. The equity premium measured relative to Treasury rates will partly be driven by the liquidity and safety of Treasuries and thus that these Treasury properties are partially responsible for the high equity premium.

Investors value the liquidity and safety attributes of Treasuries. We document this by showing that changes in Treasury supply have large effects on a variety of yield spreads.

By studying pairs of assets with similar liquidity but different safety or with similar safety but different liquidity we document that changes in Treasury supply drive both the equilibrium price of safety and the equilibrium price of liquidity. This implies that Treasuries carry both a safety and liquidity attribute.

The low yields on Treasuries due to their extreme safety and liquidity imply that Treasuries in important respects are similar to money. Evidence from quantities provide further support for this idea – the private sector expands money supply (bank deposits) when the government reduces Treasury supply.

Our results have immediate implications for asset pricing by showing that theoretical models need to incorporate priced liquidity and safety attributes in order to be empirically succeesful in understanding asset prices.

11 *Piazzesi, Monika and Martin Schneider. 2016. Payments, credit, and asset prices.

This paper studies the joint determination of payments, credit, and asset prices. The starting point is that, in modern economies, transactions occur in two layers. In the enduser layer, nonbanks — for example, households, firms and institutional investors — trade goods and securities and pay for them using payment instruments supplied by banks.

A common denominator of different payment instruments is that banks commit to accept payment instructions from their clients. As a result of those payment instructions, transactions in the enduser layer generate interbank transactions in the bank layer. Payments between customers of different banks generate interbank transfers of funds.

Interbank payments are often made with reserves, but may also be handled through various forms of short-term credit.

This paper proposes a stylized model of an economy with two layers of transactions. Endusers are households and institutional investors who must pay for some goods and securities with payment instruments — deposits or credit lines — supplied by a competitive banking sector. Banks handle endusers' payment instructions and must make some interbank payments with reserves supplied by the government. Both banks and the government incur costs of leverage that decline with the quantity and quality of available collateral, in particular securities and claims to future taxes.

The model determines asset prices, the nominal price level and agents' portfolios as a function of government policy and investor beliefs about asset payoffs. It also determines the share of resources used up as costs of leverage. An efficient payment system allocates collateral so as to minimize that share of resources and hence maximize consumption. Asset prices reflect not only cash flow expectations and uncertainty premia, but also the collateral and liquidity benefits that assets provide to endusers and banks.

We use the model think about links between securities markets and the payment system. The key properties that generate such links are illustrated by the quantity equation

$$PT = \hat{v}(D + L) \tag{1}$$

Here the total volume of transactions T includes institutional investors' securities purchases, not simply the value of goods traded. Moreover, the only medium

of exchange in the enduser layer is "inside money" — deposits D and credit lines L — provided by banks who rely on securities as collateral. Outside money — that is, reserves — is only one input to the production of inside money, albeit a special one since it not only serves as collateral, but also provides liquidity for making interbank payments.

Consider an increase in uncertainty about asset payoffs that lowers asset values. As the value of collateral that banks can purchase declines, supplying payment instruments becomes more costly. A decline in inside money D+L then puts downward pressure on the price level. At the same time, however, an increase in uncertainty also lowers institutional investors' demand for payment instruments, which has the opposite effect. The details of financial structure, including the use of payment instruments by institutional investors and the scope of netting arrangements — are thus important in order to assess the effects of asset market shocks on inflation.

We also use the model to think about recent policy shifts, with a focus on two policy tools. First, the government can trade in securities markets to change the mix of collateral available to banks. Second, the government controls the real return on reserves. The central bank sets the nominal rate on reserves. Moreover, the inflation rate is given by the growth rate of nominal government liabilities. This result follows from the quantity equation and the fact that prices are flexible. Importantly, what matters is not the growth rate of reserves, but instead the growth rate of nominal payment instruments D+L, which in turn depends on nominal collateral available to banks.

The government can select one of two policy regimes. Reserves are scarce if banks do not always have sufficient reserves to handle all interbank payments but instead turn to the short-term credit market for liquidity. The liquidity benefit of reserves then generates a spread between the short-term interest rate and the interest rate on reserves. Reserves are scarce if the real return on reserves is sufficiently low, that is, the opportunity cost of holding reserves is high. Banks then choose higher leverage to maintain a high return on equity in spite of a higher effective tax on reserves.

As long as reserves are scarce, open-market purchases of short-term debt for reserves change the collateral mix towards more liquid bank assets. In our model, open-market purchases permanently lower the real short-term interest rate. Indeed, when more liquid reserves are available, competition drives banks to produce more payment instruments, pushing the price level up. As a result, the real value of nominal collateral falls — banks become more levered and bid up the prices of all collateral including short bonds, a permanent "liquidity effect" on the real interest rate.

In the second policy regime, reserves are abundant : the quantity of reserves is sufficiently large relative to the volume of transactions that overnight borrowing is never needed. Once reserves lose their liquidity benefit, short-term loans and reserves become perfect substitutes and earn the same interest rate — the economy enters a "liquidity trap" where conventional open-market policy becomes ineffective. Reserves are abundant whenever the real return on reserves is sufficiently high, which can also happen with positive or negative interest on reserves.

The fact that payments occur in two layers has important implications for what it means to be in a "liquidity trap". The textbook view is that equality of interest rates on outside money and short bonds implies that the medium of exchange and a safe store of value become perfect substitutes for all agents. In our model, this is true only for banks who are the only investors in both reserves and short bonds. In contrast, payment instruments for endusers require costly bank leverage and never become abundant. In particular, they retain their liquidity benefit even when reserves are abundant. At the same time, collateral remains scarce in the liquidity trap, so unconventional policy that exchanges reserves for lower quality collateral can still matter by changing the collateral mix.

Which regime is better depends on the relative leverage costs of banks versus the government. If the government can borrow more cheaply than banks, then it makes sense to move to abundant reserves, as several central banks have done recently. In contrast, if the government has trouble to credibly commit to a path for nominal debt, then it is beneficial to have banks rely more on collateral other than government debt or reserves. Since the optimal system depends on the quality of collateral, it may make sense to switch between regimes over time in response to asset market events.

The availability of two separate policy tools implies that the stance of policy cannot be easily summarized by a single variable, such as the short-term nominal interest rate. For example, when reserves are scarce, the government can lower the nominal interest rate either through open-market purchases or by lowering the real return on reserves. However, the effect on real interest rates and inflation is generally different. The reason is that asset values reflect not only liquidity benefits — as in many monetary models — but also collateral benefits. Policy affects interest rates by altering both benefits.

12 *Lucas, Robert E., Jr. 2014. Liquidity : meaning, measurement, management.

The maintenance of consistently low inflation rates is an important goal that central banks can actually achieve. Financial stability is another classic responsibility of the Fed. Financial panics are the results of sudden declines in liquidity.

The main tool the Fed has for influencing the inflation rate is its ability to add to or subtract from the amount of government-issued money—I'll call it cash—in the hands of the public. This is done, of course, by selling government bonds for cash, decreasing the cash held by the public, or buying bonds with cash, and so increasing the cash outstanding. Here I take a Monetarist view that (i) open market operations affect the quantity of liquidity in the system—the money supply, if you like—and (ii) changes in this magnitude are what determine inflation rates.

These operations affect interest rates, but the assets used in monetary policy are money market instruments with yields that are not closely related to interest rates in general. These are specialized assets that are useful in cash management : interest-bearing securities that can be easily converted into non-interest-bearing cash as payments come due.

The only reason anyone ever holds cash is to be ready to make payments to someone. There is no other reason to hold this dominated asset—an asset that has no risk advantage over other assets that offer a higher return. This fact gives rise to a host of other assets that people believe can be traded for cash on short notice, on predictable terms, and without undue labor costs. These qualities define the terms liquidity and liquid asset.

Like cash, all these money market instruments have lower yields than other assets with the same risk characteristics : They command a liquidity premium.

Empirically we observe a one-for-one effects of money growth on inflation and that there is no systematic relation between real and nominal growth, but some kind of relation cannot be ruled out.

Accepting some form of the quantity theory of money, then, we have a theory of inflation. How do we apply this theory to control the inflation rate ? One possibility would be to commit to a constant rate of money growth, chosen to imply the desired inflation rate. But what if the theory is not quite accurate or becomes obsolete because of changes in banking practices ? A better, more flexible idea is inflation targeting. The central bank sets a target and if actual inflation goes above the targeted range, the bank reduces the rate of money growth. If inflation continues

above, the bank reduces the money growth rate some more and keeps taking money out of the system until inflation is brought back within the targeted range.

I have described inflation targeting as though it is implemented by adding to or reducing the money supply. As noted previously, today it is more customary to describe monetary policy as carried out by changes in short-term money market rates, not as changes in the money supply. In normal times, changes in the money supply are effected through trades in money markets ; expanding the money supply and reducing the federal funds rate are just two ways of saying the same thing.

The quantity theory that I have described is consistent with the idea that a sudden reduction in the money supply leads to deflation and reduction in spending. This happens because a sudden loss of liquidity leads people to reduce spending to rebuild a desired ratio of cash to spending flows. But central banks do not just suddenly reduce the money supply. The source of a financial crisis or panic generally lies in a loss of confidence in privately issued promises of cash, not in reductions in government supplied money.

Financial firms make huge transactions. Imagine trying to carry out trading at this volume using government-provided cash, paying no interest or nearly so, to settle accounts. The opportunity cost would be staggering. Promises to pay have substituted for payments in cash

They do this today by constructing short-term liquid assets from pieces of illiquid securities. Some of these securities were derived from mortgages. Creditors accept it only when they think they can pass it on. What we mean by a "financial crisis" is an occasion when this confidence vanishes. The effect of this situation is that financial firms that were working with minimal amounts of cash are now cash hungry. But there is only so much cash out there, so this can only mean that cash available to households and nonfinancial firms has to shrink.

In quantity-theoretic terms, the money supply available for purchases of goods and services shrinks drastically, implying deflation or losses of production or both.

13 *Teles, Pedro and Ruilin Zhou. 2005. A stable money demand : looking for the right monetary aggregate.

It is conventional to call the relationship between real money, a nominal interest rate, and a measure of economic activity a money demand relationship. A stable relationship between these variables helps answer important questions such as the

following : What is the average growth rate of money that is consistent with price stability, given the average growth of the economy and a stable nominal interest rate ? Knowledge about the response of money demand to changes in the nominal interest rate may also help quantify the welfare gains from a low average inflation rate.

Lucas (1988) shows that there is a theoretical equilibrium relationship between real money, a nominal interest rate as a measure of the opportunity cost of money, and gross domestic product (GDP) as a measure of transactions that is not exactly a money demand, but that is indeed stable. He estimates that equilibrium relationship using the monetary aggregate M1 as the measure of money with data up to 1985 and argues that there is a stable relationship between those variables with a unitary income elasticity and with a strong negative response of real balances to the nominal interest rate.

The relationship estimated by Lucas (1988) holds very well until the mid-1980s but not well at all after that. We argue that technological innovation and changes in regulatory practices in the past two decades have made other monetary aggregates as liquid as M1, so that the measure of money should be adjusted accordingly. We show that once a more appropriate measure of money is taken into consideration, the stability of money demand is recovered.

Until the end of the 1970s, the transactions demand for money was well approximated by M1. Since then, however, a series of sweeping regulatory reforms and technological developments in the banking sector have significantly changed the way banks operate and the way people use banking services and conduct transactions.

Nationwide negotiable orders of withdrawal accounts (NOWs), which are interest-bearing checking accounts classified in M1 were authorized. Also money market deposit accounts (MMDAs), interest-bearing savings accounts that can be used for transactions with some restrictions. MMDAs are classified in M2.

These two major banking reforms blurred the traditional distinction between the monetary aggregates M1 and M2 in their transactions and savings roles. The rapid development of electronic payments technology and, in particular, the growing use of credit cards and the automated clearinghouse (ACH) as means of payment, reinforced the effect of the banking reforms in slowing down the growth of M1.

The widespread adoption of retail sweep programs by depository institutions since 1994, which reclassify checking account deposits as saving deposits overnight, reduced the balances that were classified in M1 by almost half.

These fundamental changes in the regulatory environment and the transactions technology justify the use of a different measure of money after 1980. The measure MZM (money zero maturity) includes balances that can be used for transactions immediately at zero cost as a more appropriate measure of the transactions demand for money. We show that changing the monetary aggregate measure from M1 to MZM from 1980 onward preserves the long-run relationship between real money, the opportunity cost of money, and economic activity up to a constant factor.

14 *Atkeson, Andrew and Lee E. Ohanian. 2001. Are Phillips curves useful for forecasting inflation ?

A Phillips curve is an equation that relates the unemployment rate, or some other measure of aggregate economic activity, to a measure of the inflation rate. Modern specifications of Phillips curve equations relate the current rate of unemployment to future changes in the rate of inflation. These specifications are based on the idea that there is a baseline rate of unemployment at which inflation tends to remain constant. The idea is that when unemployment is below this baseline rate, inflation tends to rise over time, and when unemployment is above this rate, inflation tends to fall. The baseline unemployment rate is known as the non-accelerating inflation rate of unemployment (the NAIRU), and modern specifications based on it are known as NAIRU Phillips curves.

NAIRU Phillips curves are widely used to produce inflation forecasts, both in the academic literature on inflation forecasting and in policymaking institutions.

We examine the accuracy of three sets of NAIRU Phillips curve–based inflation forecasts. To evaluate the usefulness of Phillips curves for forecasting inflation, we compare the accuracy of these three sets of inflation forecasts at a one-year forecast horizon to that of a naive model that makes a simple prediction : at any date, the inflation rate over the coming year is expected to be the same as the inflation rate over the past year.

Our result contrasts sharply with the conventional wisdom. We find that over the last 15 years, all three sets of NAIRU Phillips curve–based inflation forecasts have been no more accurate than the forecast from our naive model.

The likelihood of accurately predicting a change in the inflation rate from these three forecasts is no better than the likelihood of accurately predicting a change based on a coin flip.

15 *Lucas, Robert E., Jr. 2000. Inflation and welfare.

This paper surveys research on the welfare cost of inflation. New estimates are provided, based on U.S. time series for 1900-94, interpreted in a variety of ways. It is estimated that the gain from reducing the annual inflation rate from 10 percent to zero is equivalent to an increase in real income of slightly less than one percent. Using aggregate evidence only, it may not be possible to estimate reliably the gains from reducing inflation further, to a rate consistent with zero nominal interest.

Since the opportunity cost of holding non-interest-bearing money is the nominal rate of interest, we would expect that the time people spend tiying to economize on cash holdings should be an increasing function of the interest rate. This observation is consistent with much evidence, and suggests that as long as interest rates are positive people could be made better off if money growth, and hence the average inflation rate and the interest rate, were reduced.

In all of the models I have reviewed, the estimated gains of reducing inflation and interest rates are positive, starting from any interest rate above, say, one tenth of one percent.

A 3 percent interest rate is about the rate that would arise in the U.S. economy under a policy of zero inflation. The optimal monetary policy, within the class of theories discussed in this paper, entails a deflation consistent with interest rates at or near zero.

16 **Diamond, Douglas W. and Philip H. Dybvig. 1983. Bank runs, deposit insurance, and liquidity.

During a bank run, depositors rush to withdraw their deposits because they expect the bank to fail. In fact, the sudden withdrawals can force the bank to liquidate many of its assets at a loss and to fail. In a panic with many bank failures, there is a disruption of the monetary system and a reduction in production.

The model we present has an explicit economic role for banks to perform, the transformation of illiquid assets into liquid liabilities. This paper gives the first explicit analysis of the demand for liquidity and the "transformation" service provided by banks. Uninsured demand deposit contracts are able to provide liquidity

but leave banks vulnerable to runs. This vulnerability occurs because there are multiple equilibria with differing levels of confidence.

Our model demonstrates three important points. First, banks issuing demand deposits can improve on a competitive market by providing better risk sharing among people who need to consume at different random times. Second, the demand deposit contract providing this improvement has an undesirable equilibrium (a bank run) in which all depositors panic and withdraw immediately, including even those who would prefer to leave their deposits in if they were not concerned about the bank failing. Third, bank runs cause real economic problems because even "healthy" banks can fail, causing the recall of loans and the termination of productive investment. In addition, our model provides a suitable framework for analysis of the devices traditionally used to stop or prevent bank runs, namely, suspension of convertibility and demand deposit insurance.

The illiquidity of assets enters our model through the economy's riskless production activity. The technology provides low levels of output per unit of input if operated for a single period but high levels of output if operated for two periods. The analysis would be the same if the asset were illiquid because of selling costs : one receives a low return if unexpectedly forced to "liquidate" early. In fact, this illiquidity is a property of the financial assets in the economy in our model, even though they are traded in competitive markets with no transaction costs. Agents will be concerned about the cost of being forced into early liquidation of these assets and will write contracts which reflect this cost. Investors face private risks which are not directly insurable because they are not publicly verifiable. Under optimal risk sharing, this private risk implies that agents have different time patterns of return in different private information states and that agents want to allocate wealth unequally across private information states. Because only the agent ever observes the private information state, it is impossible to write insurance contracts in which the payoff depends directly on private information, without an explicit mechanism for information flow. Therefore, simple competitive markets cannot provide this liquidity insurance.

Banks are able to transform illiquid assets by offering liabilities with a different, smoother pattern of returns over time than the illiquid assets offer. These contracts have multiple equilibria. If confidence is maintained, there can be efficient risk sharing, because in that equilibrium a withdrawal will indicate that a depositor should withdraw under optimal risk sharing. If agents panic, there is a bank run and incentives are distorted. In that equilibrium, everyone rushes in to withdraw their deposits before the bank gives out all of its assets.

Illiquidity of assets provides the rationale both for the existence of banks and for

their vulnerability to runs. An important property of our model of banks and bank runs is that runs are costly and reduce social welfare by interrupting production (when loans are called) and by destroying optimal risk sharing among depositors.

We analyze bank contracts that can prevent runs and examine their optimality. We show that there is a feasible contract that allows banks both to prevent runs and to provide optimal risk sharing by convert- ing illiquid assets. The contract corresponds to suspension of convertibility of deposits (to currency), a weapon banks have historically used against runs. Under other conditions, the best contract that banks can offer (roughly, the suspension-of-convertibility contract) does not achieve optimal risk sharing. However, in this more general case there is a contract which achieves the unconstrained optimum when government deposit insurance is available.

What is crucial is that deposit insurance frees the asset liquidation policy from strict dependence on the volume of withdrawals.

Our results have far-reaching policy implications, because they imply that the real damage from bank runs is primarily from the direct damage occurring when recalling loans interrupts production.

17 *Friedman, Milton and Anna J Schwartz. 1971. A monetary history of the United States, 1867-1960.

The contraction from 1929 to 1933 was by far the most severe businces-cycle contraction during the near-century of U.S. history we cover and it may well have been the most severe in the whole of U.S. history.

U.S. net national product in current prices fell by more than one-half from 1929 to 1933 ; net national product in constant prices, by more than one-third.

No other contraction before or sirice has been preceded by such a long period over which the money stock failed to rise. Monetary behavior during the contraction itself is even more striking. From the cyclical penk in August 1929 to thc cyclical trough in March 1933, the stock of money fell by over a third. This is more than triple the largest preceding declines recorded in our series.

More than one-fifth of the commercial banks in the United States holding nearly one-tenth of the volume of deposits at the beginning of the contraction suspended operations because of financial difficulties.

The 1929-33 contraction had far-reaching effects in many directions, not least on monetary institutions and academic and popular thinking about the role of monetary factors in the economy. A number of special monetary institutions were established in the course of the contraction, and the powers of the Federal Reserve System were substantially modified. The contraction was shortly followed by the enactment of federal insurance of bank deposits .

Opinion shifted almost to the extreme that "money does not matter"; that it is a passive factor which chiefly reflects the effects of other forces; and that monetary policy is of extremely limited value in promoting stability. The evidence summarized in the rest of this chapter suggests that these judgments are not valid inferences from experience.

The monetary collapse was not the inescapable consequence of other forces, but rather a largely independent factor which exerted a powerful influence on the course of events. The failure of the Fed to prevent the collapse reflected not the impotence of monetary policy but rather the particular policies followed by the monetary authorities and, in smaller degree, the particular monetary arrangements in existence.

True, as events unfolded, the decline in the stock of money and near-collapse of the banking system can be regarded as a consequence of nonmonetary forces in the United States, and monetary and nonmonetary forces in the rest of the world. Everything depends on how much is taken as given. For it is also trues, that different and feasible actions by the monetary authorities could have prevented the decline in the stock of money -indeed, could have produced almost any desired increase in the money stock. The same actions would also have eased the banking difficulties appreciably. Prevention or moderation of the decline in the stock of money, would have reduced the contraction's severity and almost certainly its duration.