FEDERAL RESERVE BANK *of* CHICAGO
FINANCIAL MARKETS GROUP

PDP 2008-6

Stationarity without Degeneracy in a Model of Commodity Money

Ricardo Cavalcanti and Daniela Puzzello

October 17, 2008

POLICY DISCUSSION PAPER SERIES

Stationarity without Degeneracy in a Model of Commodity Money*

R. de O. Cavalcanti[1] and D. Puzzello[2]
[1]EPGE/FGV, Praia de Botafogo 190, Rio de Janeiro, RJ 22250, Brazil
ricardo.cavalcanti@fgv.br
[2]Department of Economics, University of Illinois, Champaign, IL 61820, USA
dpuzzell@uiuc.edu

October 17, 2008

Abstract

We develop a model of macroeconomic heterogeneity inspired by the Kiyotaki-Wright (1989) formulation of commodity money, with the addition of linear utility and idiosyncratic shocks to savings. We consider two environments. In the *benchmark* case, the consumer in a meeting is chosen randomly. In the *auctions* case, the individual holding more money can be selected to be the consumer. We show that in both environments socially optimal trading decisions (that are individually acceptable) are stationary and solve a tractable static optimization problem. Savings decisions in the benchmark case are remarkably invariant to mean-preserving changes in the distribution of shocks. This result is overturned in the auctions case.

Keywords and Phrases: Macroeconomics with heterogeneous savings; commodity money with linear adjustments; mechanism design; auctions

JEL Classification: E00, C00

1 Introduction

One attractive but commonly overlooked feature of monetary models is their implications about heterogeneity. Contrary to an old tradition of aggrega-

*We thank Narayana Kocherlakota, Robert Molzon, Neil Wallace, Randy Wright, and seminar participants at the SED meetings in Prague and at the Money, Bank, Payments and Finance workshop at the Federal Reserve of Cleveland for useful comments and discussions.

tion, macroeconomists are increasingly studying uneven allocations of risk across the population. Uninsured risk is no news to monetary theory, provided that the essentiality of money is properly taken into account. Since the *invisible hand* fails or does not operate in a simplistic way in monetary models, consumption flows are instead organized by incomplete insurance and dispersion of individuals across asset holdings. Monetary theory is therefore in position to offer macroeconomists a coherent description of savings disparities resulting from a variety of patterns in monetary trades. This paper is concerned with tractable descriptions of monetary frictions and the consequent heterogeneities. We are particularly interested in how that description is facilitated by the property that savings decisions are heterogeneous but stationary, and how such an outcome can be derived on efficiency grounds.

In monetary theory, incomplete markets cannot be taken as a serious primitive in the field because different formulations of incompleteness have remarkably different implications for how money is used. Consensus is now building that it is necessary to start with physical-environment assumptions, like imperfect monitoring and commitment, in order to offer predictions about savings paths. But because models of money are well known for displaying multiplicity and nonstationarities, the *curse of dimensionality* associated to heterogeneity has become a serious obstacle for blending monetary exchange and macroeconomic questions. In this paper, we formulate a model of monetary exchange and use constrained efficiency to predict a rich but tractable heterogeneity. Our model owns its tractability to the extreme assumption that the only asset is commodity money, and that additions or subtractions of money holdings can be done at a linear utility cost, conditional on the realization of idiosyncratic shocks. It is essentially a version of the Kiyotaki-Wright (1989) model, with shocks to preferences and unbounded holdings. As in their model, no markets open in our model. Unlike their analysis, however, we do not need to impose stationarity, and are able to offer predictions about the distribution of money that maximizes ex-ante welfare.

The aim in this paper is to pursue a mechanism-design formulation of efficient allocations and to show that the optimum is constant, a result consistent with linear dynamics. In other fields, the idea of using linearity in order to facilitate the description of heterogeneity is not new.[1] Most papers in monetary theory make however an effort to restore the aggregative struc-

[1] Cavalcanti and Erosa (2007), for instance, use linearity in a version of the Lucas's tree economy, in which the productivity of a business depends on ownership shocks, in order to predict business turnover rates.

ture of traditional macro models, and leave potential descriptions of heterogeneity unexplored. In a notable exception, Galenianos and Kircher (2006) model a sequence of auctions that follow market interactions with possibly heterogeneous outcomes in savings of money. Although they sidestep questions about stationarity and optimality to some extent, their formulation of heterogeneity is similar in spirit to the one adopted here. In order to make that comparison easier for the reader, we have included a section explaining how an auctions setup in the context of our environment could help allocating resources in our model (more on this below). Lagos and Wright (2005) is a standard reference for a different motivation: to impose stationarity and to appeal to quasi-linearity and markets in order to eliminate the distribution of money and to evaluate inflationary policies. Shi (1997) had already pursued a model of degeneracy and policy evaluation, but did so with a coordination of individuals according to 'families', making it difficult to assess optimality in his model.[2] These models remain attractive because they can easily address policies related to the supply of fiat money, as they are usually stated, an issue that in the case of our model would certainly require future research.

We succeed in providing a description of optima that is remarkably simple. Instead of setting up individual choices in sequence spaces or constructing value functions, we propose the analysis of a simple static problem. We then show that the problem defines an *upper bound* on average welfare of implementable allocations, and that its solution is actually implementable as a constant sequence of consumption and savings decisions for the whole economy. Because the optimum is shown to have low dimensionality, we can ask what gives shape to the distribution of money, and pursue an answer in the context of both the *benchmark* Kiyotaki-Wright formulation, and what we call an *auctions* variation of the environment. In the benchmark environment, the consumer in a meeting is chosen as the realization of a random variable, as usual, while in the auction environment we let the planner pick who is the consumer between two candidates in a meeting (the other must be the producer). We find that the optimum has the individual with largest holdings of money being the consumer because such a choice implies the weakest constrain on average utility. The tractability of this auction environment is evidence that the *upper-bound argument* can be generalized to

[2] One concern in models in which individuals can commit to family (or another entity such as government) plans is that gift-giving would render money inessential. Even when individuals are anonymous (i.e., their identities are not observable) but meet in large groups, special assumptions are needed to rule out certain trigger strategies. See Aliprantis et al. (2007).

more applications in macroeconomics.

The contrast between the two environments sheds light on what the distribution of money represents in this linear context. We learn quite a sharp lesson about the benchmark case. Because money is costly to acquire, the optimal allocation has individuals economizing on money holdings in a way that aligns private and social returns. In particular, trade takes place in pairwise meetings with all surplus going to the consumer, and all money holdings going to the producer. Due to linearity, after-trade holdings are valued according to an average marginal utility, which is the mean of the shock distribution. At the stage of making savings decisions (before trade), individuals know that holdings of producers are irrelevant and need not predict the distribution of money (it suffices to know what the average marginal utility is in case they are called to produce). The conclusion is that *individual* savings decisions are invariant to changes in the distribution of shocks that preserve the mean. Consequently, the distribution of money can be computed residually in the benchmark case.

The picture is different in the auctions environment, although previous findings about stationarity apply. Society economizes resources by having individuals with large holdings being the consumers. Therefore, the return of money is affected by the way holdings are distributed. Savings decisions now equilibrate two forces. It equalizes intertemporal costs and returns and, at the same time, generates the distribution of money used to compute money returns. The analysis is still easily accessible. We are able to avoid a demanding fixed-point treatment in the space of distributions. We tackle instead a first-order differential equation that is necessary and sufficient to characterize optimal savings in some class. We also offer, in our conclusion, a discussion of how the findings square with a recent literature on money and auctions. In that literature, the distribution of money reflects *indifference* among directed-search opportunities, not *optimality* as in our setup.

2 The benchmark environment

Our model is a version of Kiyotaky-Wright [8] in which only one good is durable. This good, interpreted as commodity money or capital, can be consumed and produced by all individuals in the population. We refer to commodity money as just 'money' for simplicity.

Time is discrete and the horizon is infinite. The economy is populated by a continuum of individuals symmetrically divided according to an integer number of types N, where $N \geq 3$. There are N specialized goods per date and one common good called money. Specialized goods are assumed divisible and perishable. Money is assumed divisible and storable. Each period is divided into two stages, and money can only be produced or consumed in the second stage. In the first stage, people meet according to random, pairwise meetings, and consumption of (specialized) goods by one person can only be provided by the meeting partner, and only if a coincidence of wants occurs. More specifically, type j enjoys good j but can only produce good $j+1$, modulo N. Since $N > 2$, no double-coincidence in which two individuals consume can occur. In the second stage, each individual in isolation consumes or produces money, according to an individual-specific production function, to be defined below. Money can also be used in future dates to pay for consumption of specialized goods.

Utility is separable across stages. If first-stage $x \in \mathbb{R}_+$ is produced by type $j-1$ and consumed by type j, in what we call a single-coincidence meeting, then the period utility for the consumer is $u(x)$, and the period utility for the producer is $-v(x)$, independently of j. In the second stage, each individual is hit by an idiosyncratic preference/productivity shock s distributed according to the measure λ. The shock is distributed independently and identically across individuals and over time. Production of money at the second stage must be nonnegative, and if an individual in state s produces $m \in \mathbb{R}_+$ and consumes $c \in \mathbb{R}_+$ units of money then a corresponding utility flow $s(c-m)$ takes place. Individuals discount future utilities according to the common discount factor β, where $\beta \in (0,1)$, but there is no discounting across stages of the same date. The functions u and v are defined on $\mathbb{R}_+$ and assumed increasing, continuous and differentiable. In addition, u is strictly concave, v is convex, $u-v$ is bounded from above, and $u'(0) = +\infty$. In order to have a compact savings problem we impose a lower bound to s of the form $s > \beta \int s' d\lambda(s')$. For ease of exposition, we also find convenient to choose normalizing units so that $u(0) = v(0) = 0$ and $\int s d\lambda = 1$. Hence λ is assumed to have support in the Borel subsets of $S \equiv (\beta, +\infty)$.

People cannot commit to future actions and their personal histories are private. The only assets are holdings of money. We assume that money holdings and types are observable by participants in a meeting.

We also considered another economy in a second part of this paper. In that economy, there are meetings in which both individuals can produce, yet only one can be the producer, and the other one must be the consumer. We shall investigate whether an allocation resembling an auction over holdings of money can implement the optimum. Because the description of allocations in this second environment is a straightforward extension of the benchmark environment, we proceed with a discussion of implementability for the benchmark environment only, and then make the necessary qualifications in the second part of this paper.

2.1 Allocations

An allocation is a description of what happens in all stages and dates. We build on [5] and define allocations as trade and savings plans for all contingencies without imposing stationarity, but our application turns out to be much simpler, and the proof for the existence of an optimum is done by construction. As a result, we do not need to define the planner's dynamic programming problem and can avoid much the questions raised in [5] about compactness and admissibility of the state space.

We assume that the economy starts at the second stage of date zero, and with zero money endowments. Part of the planner's problem is to choose a sequence of Borelean measures $\{\mu_t\}_{t=1}^{\infty}$ describing the distribution of money holdings in $\mathbb{R}_+$ at the start of all dates.

In order to fix ideas, let us take date t and assume that the initial distribution of money, μ_t, has already been chosen. An allocation plan (for date t) is a trade plan for the first stage, (g, h), and a savings plan for the second stage, f. More formally, the function $g : \mathbb{R}_+^2 \to \mathbb{R}_+$ determines the quantity of output produced by producers for each kind of single-coincidence meeting (m, n), where the consumer is type j and has m units of money, and the producer is type $j-1$ (modulo N) and has n units of money. The function $h : \mathbb{R}_+^2 \to \mathbb{R}_+$ determines the corresponding after-trade holdings of the consumer in these single-coincidence meetings. As a result, in meeting (m, n), the after-trade holdings of the producer is $n + m - h(m, n)$. The function $f : S \times \mathbb{R}_+ \to \mathbb{R}_+$ describes final holdings for an individual at the end of the second stage, and is defined over pairs (s, m), where $s \in S$ is the productivity state at the second stage and m is the holdings of money at the beginning of the second stage.

It should be clear that an allocation plan fully describes what happens in a date. In the first stage, trade takes place in single-coincidence meetings according to (g, h). An individual starting the second stage with state (s, m) and leaving with holdings $m' = f(s, m)$ has consumed $c = \max\{m - m', 0\}$ and enjoyed utility $s(m - m')$. A plan is said feasible if $h(m, n) \le m$ for all $(m, n) \in \mathbb{R}_+^2$.

We restrict attention to allocation plans that are continuous and thus measurable functions. Similar arguments to those applied by [5] would then show that an allocation plan maps a Borelean measure μ_t into another Borelean measure μ_{t+1} in the obvious way. In particular, at $t = 0$, the economy starts at the second stage without money endowments, and a date-0 savings plan f_0, together with λ, generates μ_1. More generally, we say that the sequence $\{f_{t-1}, h_t\}_{t=1}^{\infty}$ generates $\{\mu_t\}_{t=1}^{\infty}$. An allocation, and thus a complete description of what happens in all meetings, is a sequence of feasible plans $\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$. We also include $\{\mu_t\}_{t=1}^{\infty}$ among the objects associated to an allocation whenever it is clear from the context that $\{\mu_t\}_{t=1}^{\infty}$ is implied by a particular sequence $\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$.

2.2 Implementability

When an allocation is fixed, a game is formed with individuals in each stage choosing an agree/disagree strategy at each contingency in response to the plan prescribed by the allocation. Hence we only allow individual defections. There are also two implications of sequential individual-rationality: individuals who disagree at the first stage are allowed to preserve any holdings brought to a meeting; and individuals who disagree at the second stage, are free to deviate from the plan (f) and choose different holdings $m' \in \mathbb{R}_+$ as the state for the next date. An allocation is implementable if it is feasible and if *agree* in all contingencies is a subgame perfect equilibrium.

2.3 The welfare criteria

The planner's problem is to find an implementable allocation that maximizes average discounted utility. The goal is to find a savings and goods exchange path that achieves the highest average discounted utility under implementability.

We express average utility by parts with the help of some simplifying notation. Let us use $u \circ g$ and $v \circ g$ to denote utility flows, for each (m, n), when the output is $g(m, n)$. With respect to the law of motion of holdings, $f \circ h$ in general, and $(f \circ h)(s, m, n)$ in particular (which is short for $f(s, h(m, n))$), describe final holdings of an individual after being the consumer in meeting (m, n) and drawing productivity s. Using also $\tilde{h}(m, n) \equiv n + m - h(m, n)$, with the composition $f \circ \tilde{h}$ it is understood that the individual was the producer. For completeness, we let the function i denote the projection with respect to the first coordinate of pairs (m, n), so that $i(m, n) = m$ and, hence, $(f \circ i)(s, m, n)$ describes the law of motion for those in no-coincidence meetings. Finally, in our application, double integration with respect to m and n can be written in a compact form as integration with respect to the product measure τ, where $\tau(A \times B)$ is defined as $\mu(A)\mu(B)$ for cartesian products of Borel sets A and B in $\mathbb{R}_+$.

Let us consider the consequences of allocation plan (f, g, h) for average utility at some date in which μ, and thus τ, are fixed. In single-coincidence meetings, the sum of the flow of utilities accruing to the consumer and producer in meeting (m, n) is $u(g(m, n)) - v(g(m, n))$. Hence, using a compact notation, we write the average over m and n as

$$\int (u \circ g - v \circ g) d\tau.$$

The average utility, in the second stage, for those who were consumers in those meetings is

$$\int \int s(h - f \circ h) d\tau d\lambda$$

That average for those who were producers is

$$\int \int s(\tilde{h} - f \circ \tilde{h}) d\tau d\lambda,$$

while the average utility for those experiencing no-coincidence meetings is

$$\int \int s(i - f \circ i) d\tau d\lambda.$$

An individual finds a single coincidence meeting as consumer with probability $\frac{1}{N}$, as producer with probability $\frac{1}{N}$, and none of the above with probability $1 - \frac{2}{N}$. Since $\int (h + \tilde{h}) d\tau = 2 \int i d\tau$ then

$$\int \int \left(\frac{1}{N} sh + \frac{1}{N} s\tilde{h} + (1 - \frac{2}{N}) si \right) d\lambda d\tau = \int \int si d\tau d\lambda.$$

Thus, second-stage average utility is

$$\int \int s \left(i - \left(\frac{1}{N} (f \circ h + f \circ \tilde{h}) + (1 - \frac{2}{N}) f \circ i \right) \right) d\tau d\lambda.$$

The objective to be maximized in the planner's problem is the date-0 discounted sum of average utility, or W, defined as

$$W(\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}) = -\int s f_0(s, 0) d\lambda + \sum_{t=1}^{\infty} \beta^t \int \int A_t d\tau_t d\lambda,$$

where the term A_t inside the integral is

$$si + \frac{1}{N}(u \circ g_t - v \circ g_t - s f_t \circ h_t - s f_t \circ \tilde{h}_t) - (1 - \frac{2}{N}) s f_t \circ i.$$

In the right-hand side of the equation defining W, the first term is the social cost at date $t = 0$ of savings decisions leading to the initial distribution of money, and the second is the present value of stage-1 and stage-2 utilities for $t \geq 1$.

Remark 1 *If, for all t, f_t is constant in its second argument (varies only with s) then consumption at $t+1$ of $f_t(s,\cdot)$ units of commodity at marginal utility s' gives average utility $\int s' f_t(s,\cdot)d\lambda(s') = f_t(s,\cdot)$ since, as assumed, $\int s' d\lambda(s') = 1$. Integrating now $f_t(s,\cdot)$ over s, then average utility at $t+1$ from consumption of time-t investment can be written in simple terms as $\int s f_t(s,\cdot)d\lambda$, so that the expression for W becomes*

$$-\int (s-\beta)f_0(s,0)d\lambda + \sum_{t=1}^{\infty} \beta^t \left(\frac{1}{N} \int (u-v)\circ g_t d\tau_t - \int (s-\beta) f_t(s,\cdot)d\lambda \right).$$

Notice also that by treating $u \circ g_t - v \circ g_t$ as identically equal to zero at $t = 0$, this expression for W can be further simplified to

$$\sum_{t=0}^{\infty} \beta^t \left(-\int (s-\beta) f_t(s,\cdot)d\lambda + \frac{\beta}{N} \int (u-v)\circ g_{t+1} d\tau_{t+1} \right).$$

When f_t is constant in its second argument, τ_{t+1} is uniquely determined by λ and f_t for all t. Thus, in this case, W corresponds to a discounted sum of independent terms.

3 Optimality

We shall construct a constant allocation (f^*, g^*, h^*) and show that it is optimal. All three functions in the constructed allocation are constant in the second coordinate, so that output and after-trade holdings of consumers do not depend on holdings of producers, and savings do not depend on stage-1 outcomes. We shall first present (f^*, g^*, h^*) as continuous solutions of particular problems, assuming that such solutions exist, and then show, in a lemma below, that the allocation is actually well defined. We then finish the section with the statement about optimality.

The construction proceeds in two steps. First, given some money holdings m of the buyer, the optimal exchange of goods in a single-coincidence meeting involves the maximization of the joint surplus under the constraint that there is enough money for the producer to be compensated for his disutility (see P1). Second, given this optimal trading decision, the optimal savings decision m' given some shock s is determined. This choice is pinned down by trading off the benefit with the cost of money (see P2). More details are provided next.

Consider the sum of stage-1 utility flows in a single coincidence meeting in which the consumer has m and the producer has n units of money. Consider now the problem of maximizing this over choices of output $g(m,n)$ restricted to $v(g(m,n)) \leq m$. Let now $g^*(m,n)$ be a continuous, constant-in-n, solution to

$$\max_y \{u(y) - v(y) : v(y) \leq m\} . \tag{P1}$$

We construct h^* as the slack in the constraint of the problem defining g^*, that is, $h^*(m,n) = m - v(g^*(m,n))$ for all m and n. We next consider a one-period savings problem, for a given realization $s > 0$, in order to define $f^*(s,m)$ as a continuous solution to

$$\max_{m'} \left\{ -(s-\beta)m' + \frac{\beta}{N}\left[u(g^*(m',n)) - v(g^*(m',n))\right] \right\} . \tag{P2}$$

In P2 the values of m and n are immaterial. Notice that the interpretation of P2 as a 'savings' problem applies in the sense that a choice m' implies forgone present utility sm', a possible future expected flow $\frac{1}{N}(u \circ g^* - v \circ g^*)$, and a residual expected future utility $m' \int s' d\lambda(s')$. Applying then β to discount the future and using the normalization $\int s' d\lambda(s') = 1$ yields the objective in P2. Clearly, P2 suggests that savings decisions are different depending on the costs of savings s that agents face.

Lemma 2 *There exists a continuous and bounded solution g^* to P1 so that (f^*, g^*, h^*) are well defined and constant with respect to their second coordinates.*

Proof. The objective in P1 is a continuous and concave function bounded from above. Existence follows because the constraint set is compact for $s \in S$, that is, for $s > \beta$. Uniqueness and continuity of g^* for m in $[0, v(\bar{y}))$, where $\bar{y} = \arg\max_y\{u(y) - v(y)\}$, follows from the strict concavity of $u - v$ and a straightforward application of the theorem of the maximum. For $m \geq v(\bar{y})$, continuity requires that $g^*(m,n) = \bar{y}$. Hence h^* is also continuous. That $g^*(m,n)$ and $h^*(m,n)$ are bounded and constant in n is trivial. Since g^* is bounded and continuous then f^* is continuous and well defined.[3] ■

[3] The optimization problem P1 is an optimization problem in parametric form, where the parameter is given by the buyer's money holdings. Depending on the values of the parameter m, the set of maximizers is a correspondence. In this paper, without loss of generality we focus on a continuous selection.

Proposition 3 *(i) The constant allocation (f^*, g^*, h^*) is implementable, (ii) W is bounded from above by $W(f^*, g^*, h^*)$ on the set of implementable allocations and, therefore, (f^*, g^*, h^*) is optimal.*

Proof. We start discussing implications of second-stage deviations for implementability. Let an arbitrary implementable allocation $\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$ be fixed, and consider a given date t. Since individuals can freely deviate from f_t at the second stage, and such deviations produce a linear payoff, the continuation utilities associated to after-trading in the first-stage must be linear in money holdings. As a result, producers in a single-coincidence meeting (m, n) say agree if and only if $-v(g_t(m,n)) + \tilde{h}_t(m,n) \int s' d\lambda \geq n \int s' d\lambda$, and consumers say agree if and only if $u(g_t(m,n)) + h_t(m,n) \int s' d\lambda \geq m \int s' d\lambda$. Since $\int s' d\lambda = 1$, the participation constraints implied by linearity and individual rationality are equivalent to

$$u(g_t(m,n)) \geq m - h_t(m,n) \geq v(g_t(m,n)). \tag{1}$$

Moreover, since continuation utilities are linear in after-trade holdings, then linearity of the payoff on savings decisions imply that only one-period deviations needed to be considered at a time. In particular, the payoff R_{t-1} at $t-1$ from a stage-2 deviation to m', when the shock is s and the distribution at t is μ_t, can be written as

$$\begin{aligned} &R_{t-1}(m', s, \mu_t) = -sm' + (1 - \frac{2}{N})\beta m' + \\ &\frac{\beta}{N} \int \left[u(g_t(m',n)) + h_t(m',n) - v(g_t(n,m')) + \tilde{h}_t(n,m') \right] d\mu_t(n). \end{aligned} \tag{2}$$

Hence, $\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$ is implementable if and only if (1) holds and $f_t(s, \cdot)$ maximizes $R_t(m, s, \mu_{t+1})$ in m at all dates.

We now argue that (f^*, g^*, h^*) is implementable. It is straightforward to verify that. Because g^* solves problem P1, then $y = g^*(m,n)$ and $x = h^*(m,n)$ solve

$$\max_{(x,y)} \{u(y) - v(y) : u(y) \geq m - x \geq v(y)\}.$$

Hence g^* and h^* satisfy the participation constraints (1), and individuals say agree to (f^*, g^*, h^*) in all first-stage meetings. We now claim that individuals agree with f^* at the second stage of all dates. This follows because f^* is constructed as the maximizer of problem P2, and because (2) coincides with the objective in P2 on $[0, \bar{y}]$ (defined in the proof of Lemma 1) when $g_{t+1} = g^*$

and $h_{t+1} = h^*$ (because n drops out in this case from (2) the value of μ_{t+1} becomes irrelevant). Hence individuals also say agree to (f^*, g^*, h^*) at the second stage in all dates and the proof to the first part of the proposition is complete.

We now show that $W(f^*, g^*, h^*)$ is an upper bound on welfare of implementable allocations. Let us again fix an alternative, implementable allocation $\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$ and assume that $\{\tau_t\}_{t=1}^{\infty}$ is its corresponding sequence of meetings distributions. Let $t \geq 0$ be arbitrary. Because (g^*, h^*) was constructed so as to maximize $u - v$ subject to participation constraints in any meeting (m, n), it follows that

$$(u - v) \circ g^* \geq (u - v) \circ g_{t+1}.$$

Integrating the objective in problem P2 with respect to λ implies that f^* is such that, for any m and n,

$$\begin{aligned}
& \int \left\{ -(s-\beta) f^*(s,\cdot) + \frac{\beta}{N}(u-v) \circ g^*(f^*(s,\cdot),\cdot) \right\} d\lambda(s) \\
\geq \; & \int \left\{ -(s-\beta) f_t(s,m) + \frac{\beta}{N}(u-v) \circ g^*(f_t(s,m),\cdot) \right\} d\lambda(s) \\
\geq \; & \int \left\{ -(s-\beta) f_t(s,m) + \frac{\beta}{N}(u-v) \circ g_{t+1}(f_t(s,m),n) \right\} d\lambda(s).
\end{aligned}$$

Changing now variables in both the first and the third integrals yields

$$\begin{aligned}
& -\int (s-\beta) f^*(s,\cdot) d\lambda + \frac{\beta}{N} \int \left[u(g^*(m',\cdot)) - v(g^*(m',\cdot)) \right] d\lambda f^{*-1}(m') \\
\geq \; & -\int\int (s-\beta) f_t(s,m) d\lambda d\omega_t(m) + \frac{\beta}{N} \int (u \circ g_{t+1} - v \circ g_{t+1}) d\tau_{t+1},
\end{aligned}$$

where ω_t is the distribution of after-trade holdings associated with $\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$, and λf^{*-1} is the measure of end-of-stage-2 holdings induced by f^* [that is, if A is Borel subset of $\mathbb{R}_+$ then $\lambda f^{*-1}(A) = \lambda(f^{*-1}(A))$]. For this last step, we use the fact that ω_t and f_t must generate τ_{t+1}. It is now trivial to verify that this inequality implies

$$W(f^*, g^*, h^*) \geq W(\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}).$$

The proof of the proposition is thus complete. ■

Remark 4 *This result shows that the simple problem P2 gives an upper bound on average discounted welfare for any allocation that is individually* *rational. Since it can be implemented in an individually rational way, the solution to P2 provides an optimal plan.*[4]

The constraint on the first part of the problem P1 relies on the expected value of money, thus the optimum is invariant to mean preserving changes in the distribution of shocks. This is no longer the case in the environment with auctions considered next.

4 The environment with auctions

We change the environment in order to allow for a selection of consumers in some meetings.

4.1 Mixing standard and auction meetings

For each meeting (m, n) there is now a realization of a Bernoulli shock. Its distribution is *iid* over meetings and dates. With probability $1-\pi$ consumption and production must take place as in the benchmark environment; that is, there is a single coincidence only if the consumer is type j and the producer is type $j-1$, for some j. With probability π, however, both individuals can produce the good that the meeting partner likes, but only one of the traders can be the producer.

Our goal is to show that a modification of the upper-bound argument used to construct the optimum in the benchmark case can also be used to describe the optimum in this new environment. As we shall see, an interesting aspect of the optimum is that it is desirable to choose the individual holding the largest quantity of money to be the consumer, as if that choice is the result of a first-price auction (i.e., the probability of being a buyer or a seller is endogenous). The reader can easily verify that the essence of this optimum would be preserved in extensions where the number of people in a meeting is larger than two, provided that only one individual can be the consumer, and only one of the remaining participants in the meeting can be the producer.

Let us call this new kind of meeting, an event that happens with probability π, an *auction meeting*, and call the other realization, that happens with

[4] Another implication of this result is that optimal trading exchanges involve the buyer take-it-or-leave-it offer. In other words, we show that the buyer take-it-or-leave-it trading rule is optimal among *all* the incentive compatible trading rules (including also the ones arising under Nash bargaining).

probability $1-\pi$, a *standard meeting*. We restrict the set of allocations as follows. An allocation is now $\{e_t, f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$, where the last three sequences can be given the same interpretation as before, while $e_t : \mathbb{R}_+^2 \to \{0,1\}$ defines a selection of who is the consumer in auction meetings. If type j holds m, type k holds n, and the meeting is an auction, then $e_t(m,n) = 1$ means the consumer is the one holding the largest amount (if the holdings are not equal) and $e_t(m,n) = 0$ means the consumer is the one holding the least amount. (For completeness, one can assume that the consumer is chosen according to a randomization device with probability $\frac{1}{2}$ when $m = n$).

If $m \vee n$ denotes $\max\{m,n\}$ and $m \wedge n$ denotes $\min\{m,n\}$ then, after the consumer is selected in an auction meeting, output is

$$q_t(m,n) = e_t(m,n)g_t(m \vee n, m \wedge n) + [1 - e_t(m,n)]g_t(m \wedge n, m \vee n)$$

while the after-trade holdings of the consumer is

$$p_t(m,n) = e_t(m,n)h_t(m \vee n, m \wedge n) + [1 - e_t(m,n)]h_t(m \wedge n, m \vee n).$$

The set of allocations is restricted because we are not allowing (g,h) to vary across auctions and standard meetings, but this restriction imposes no loss of generality as it follows that the optimum features $e_t = 1$ in all auction meetings. Before we present the argument, we conclude the presentation of the environment with the definition of implementability and the welfare criteria.

If $\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$ is implementable in the benchmark environment, then $\{e_t, f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$ is implementable in the new environment because saying agree or disagree to e_t presents no new participation constraints.

Hence $\{e_t, f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$ is implementable if and only if $\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$ is implementable in the benchmark environment. The expression for average utility $W(\{e_t, f_{t-1}, g_t, h_t\}_{t=1}^{\infty})$ is the same as that for $W(\{f_{t-1}, g_t, h_t\}_{t=1}^{\infty})$, with the exception that the term A_t inside the integral is now extended to

$$(1-\pi)A_t + \pi(u \circ q_t - v \circ q_t + 2si - sf_t \circ p_t - sf_t \circ \tilde{p}_t),$$

with the understanding that $\tilde{p}_t(m,n) = n + m - p_t(m,n)$.

We assume that λ has a density bounded away from zero on the interval $\bar{S} \equiv (\beta, \bar{s})$, for some $\bar{s} > \beta$.

4.2 The upper bound

As before, we first construct a particular candidate, and show later that it is in fact an optimal allocation. In our construction, the candidate is stationary (constant in t). On the one hand, in all auction meetings, $e = 1$. On the other hand, in both standard and auction meetings, trade takes place according to g^* and h^*, as in the exchange scheme for the benchmark allocation (but now, in auction meetings $g^*(m,n)$ and $h^*(m,n)$ are applied to $m \geq n$). The novelty is that the optimum savings function, now denoted f^a, is chosen among possibly many solutions to a fixed-point problem.

We find it convenient to limit ourselves to smooth and monotone solutions as follows. A function $f : \bar{S} \times \mathbb{R}_+ \to \mathbb{R}_+$ is a candidate optima if it is differentiable, strictly decreasing in the first argument, constant in the second, and solves moreover

$$\max_m \left\{ -(s-\beta)m + (1-\pi)\frac{1}{N}R(m) + \pi F(m)R(m) \right\} \tag{P3}$$

where, for all $m \geq 0$,

$$R(m) = \beta \max_y \{u(y) - v(y) : v(y) \leq m\}$$

and

$$F(m) = \lambda(\{s \in \bar{S} : f(s,\cdot) \leq m\}). \tag{FP}$$

In P3, the function R is a short representation for $\beta(u \circ g^* - v \circ g^*)$. It follows that when $\pi = 0$, the unique candidate is $f = f^*$, and thus $f^a = f^*$. For $\pi > 0$, however, problem P3 differs from P2 because now the distribution of money is relevant. For each candidate f, in allocations of the kind $(e, f, g, h) = (1, f, g^*, h^*)$ a savings level m is payoff-relevant in auction meetings only when the producer has $n \leq m$, an event taking place with probability $F(m)$.

4.3 Existence and uniqueness

In this subsection we present sufficient conditions for the existence of a unique solution to P3. We examine optimality in the next subsection.

Lemma 5 *Let D denote the cdf of λ and let f^{-1} denote the inverse of f (with respect to s). Then f is a solution to P3 if and only if $D \circ f^{-1} = x$ for $x : I \to [0,1]$ solving*

$$x' = \left(\frac{1-\pi}{\pi N} + 1 - x\right)\frac{R'}{R} - \frac{1}{\pi R}\left(D^{-1}(x) - \beta\right), \tag{DE}$$

together with the auxiliary condition

$$x(m_0) = 1\,, \tag{IC}$$

where $I = [m_0, m^*)$ is defined by $R'(m_0) = (\bar{s}-\beta)N/(1-\pi)$ and $R'(m^*) = 0$.

Proof. Recall that R' is strictly decreasing, with $R'(0) = +\infty$ and $R'(m^*) = 0$. Moreover, if f is continuous then F is differentiable and the solution $m = f(s, \cdot)$ for each s must be interior and satisfy the tangency condition

$$s - \beta = (1 - \pi)\frac{1}{N}R'(m) + \pi[R'(m)F(m) + R(m)F'(m)].$$

Since f is decreasing, $F(f(\bar{s}, \cdot)) = 0$. Thus, taking into account the definition of m_0, if $f(\bar{s}, \cdot) \neq m_0$ then

$$-(\bar{s} - \beta)m_0 + (1 - \pi)\frac{1}{N}R(m_0) > -(s - \beta)f(\bar{s}, \cdot) + (1 - \pi)\frac{1}{N}R(f(\bar{s}, \cdot)),$$

contradicting that f is a solution to P3. Hence $f(\bar{s}, \cdot) = m_0$ and a similar argument shows that $f(\beta^+, \cdot) = m^*$. Let now $h : I \to (\beta, \bar{s}]$ be the inverse of f with respect to the first coordinate. Since the FP condition is equivalent to $F(m) = 1 - D[h(m)]$, where D is the *cdf* of λ, then $F'(m) = -d(h(m))h'(m)$ for m in the image of f, where d is the density of λ. Changing variables once again with $x : (m_0, m^*) \to (0, 1)$, $x(m) = D(h(m))$ then, because $s = D^{-1}(x(m))$ and $F' = -x'$, this first order condition can be written as the differential equation (DE), with the understanding that x, x', R, and R' are functions of m. Because $f(\bar{s}, \cdot) = m_0$, then an initial condition for (DE) is (IC). ■

Lemma 6 *The initial value problem (DE)-(IC) has at most one solution. If the distribution of shocks is uniform, or $\bar{s}$ is sufficiently low, then it has a solution.*

Proof. As it is standard, in order to examine existence and uniqueness of solutions to (DE-IC) on the interval I, we let

$$H(t, r) = \left(\frac{1 - \pi}{\pi N} + 1 - r\right)\frac{R'(t)}{R(t)} - \frac{D^{-1}(r) - \beta}{\pi R(t)}$$

denote the function defined by the right-hand side of (DE) with $r = x$ and $t = m$, with the understanding that R and R' are continuous and bounded functions on I (since $\pi < 1$ implies $m_0 > 0$ and $R'(m_0) < \infty$). If λ is uniform then D is linear, and so H is also linear in r. As a result, (DE) defines a linear ordinary differential equation with nonconstant coefficients. A basic result in the theory of linear ordinary differential equations is that initial value problems are uniquely solvable, and solutions are defined on all of I.

Even if D is not linear, because H and $\frac{\partial H}{\partial r}$ are bounded (since the density of λ is assumed bounded away from zero), then another basic result on the theory of first-order differential equations is that the initial value problem (DE-IC) has at most one solution (see Theorem 2 in Brauer and Nohel, 1986, p. 400). In this case, existence on some interval can be demonstrated according to details of the bounds on H according to the Picard-Lindelöf theorem (see Theorem 1 in Brauer and Nohel, 1986, p. 389). To do that, we have to find bounds of the right-hand side of (DE) so that the interval around m_0 for which the theorem applies includes (m_0, m^*). We state below the theorem for completeness.

Theorem. Assume that M is a bound for $|H|$, that $a = m^* - m_0$, and that $b = 1$. Suppose that H and $\frac{\partial H}{\partial r}$ are continuous and bounded in the rectangle $\{(t,r) : t - m_0 < a, 1 - r < b\}$ with $|H| \leq M$. Let $\alpha = \min\{a, \frac{b}{M}\}$. Then the successive approximations ϕ_j given by

$$\begin{aligned} \phi_0(t) &= 1, \\ \phi_{j+1}(t) &= 1 + \int_{m_0}^{t} H(m, \phi_j(m))dm \qquad (j = 0, 1, 2, ...) \end{aligned}$$

converge (uniformly) on the interval $\{t : t - m_0 < \alpha\}$ to a solution ϕ of DE that satisfies initial condition IC.

We now provide a bound M for H such that α implies $(m_0, m^*) \subset \{t : t - m_0 < \alpha\}$. On the one hand, since $D^{-1} \geq \beta$ and $r \geq 0$, then

$$H(t,r) \leq \left(\frac{1-\pi}{\pi N} + 1\right)\frac{R'(t)}{R(t)} \leq \left(\frac{1-\pi}{\pi N} + 1\right)\frac{R'(m_0)}{R(m_0)}.$$

On the other hand, since $r \leq 1$,

$$H(t,r) \geq \frac{1-\pi}{\pi N}\frac{R'(t)}{R(t)} - \frac{D^{-1}(r) - \beta}{\pi R(t)} \geq -\frac{\bar{s} - \beta}{\pi R(m_0)}.$$

Hence

$$|H| \leq \frac{1}{\pi R(m_0)} \max\left\{\left(\pi + (1-\pi)\frac{1}{N}\right)R'(m_0), \bar{s} - \beta\right\},$$

and since, by construction, $\bar{s} - \beta = (1-\pi)\frac{1}{N}R'(m_0)$, then $|H| \leq M$, where

$$M = \left(\pi + (1-\pi)\frac{1}{N}\right)\frac{R'(m_0)}{\pi R(m_0)}.$$

Since

$$\alpha = \min\{a, \frac{b}{M}\} = \min\{m^* - m_0, \frac{1}{M}\},$$

in order to have $\alpha \geq m^* - m_0$, as desired, it suffices to have m_0 sufficiently large so that

$$\left(\pi + (1-\pi)\frac{1}{N}\right)\frac{R'(m_0)}{\pi R(m_0)} \leq \frac{1}{m^* - m_0}.$$

Making now explicit the dependence of R on parameters, we find that

$$\frac{R'(m_0)}{R(m_0)} = \frac{1}{u(y_0) - v(y_0)}\left(\frac{u'(y_0)}{v'(y_0)} - 1\right)$$

and

$$\bar{s} - \beta = (1-\pi)\frac{\beta}{N}\left(\frac{u'(y_0)}{v'(y_0)} - 1\right),$$

where $y_0 = v^{-1}(m_0)$. Hence, if $\bar{s} \to \beta$ then $m_0 \to m^*$. Thus $\alpha = m^* - m_0$ if $\bar{s} - \beta$ is sufficiently low. The proof is now complete. ■

4.4 Optimality with auctions

We compare solutions to P3 in terms of the average payoff ω, defined for each f as

$$\omega = \int \left\{ -(s-\beta) f(s,\cdot) + \frac{1-\pi}{N} R(f(s,\cdot)) + \pi F(f(s,\cdot)) R(f(s,\cdot)) \right\} d\lambda(s).$$

Proposition 7 *(i) If f solves P3 then $(1, f, g^*, h^*)$ is implementable. (ii) If $\{e_t, f_{t-1}, g_t, h_t\}_{t=1}^{\infty}$ is implementable and $\bar{\omega}$ is an upper bound on the average payoff in P3 then $W(\{e_t, f_{t-1}, g_t, h_t\}_{t=1}^{\infty}) \leq \bar{\omega}/(1-\beta)$.*

Proof. The same reasoning of the proof of Proposition 1 applies to statements (i) and (ii). Let us suppose f solves P3 and let us fix allocation $(1, f, g^*, h^*)$. Since individuals take the distribution of money as given (pdf F defined in P3), individual rationality is equivalent to having f maximize the right-hand side of P3, as assumed. Thus $(1, f, g^*, h^*)$ is implementable.

Regarding part (ii), it is straightforward to show the producer constrained is weakened by increases in holdings of the consumer. More formally, that $(u - v) \circ g^*(m, n)$ is weakly increasing in m and constant in n. Hence, the social payoff $u - v$ is bounded from above by the choice $e = 1$ and $g = g^*$. As in the proof of Proposition 1, integrating individuals' payoffs resulting with respect to λ yields again an aggregate payoff ω that coincides with the average objective in P3. As a result, $(1-\beta)W(\{e_t, f_{t-1}, g_t, h_t\}_{t=1}^{\infty})$ is bounded above by any $\bar{\omega}$ bounding the average payoff of the set of solutions to P3. ■

Remark 8 *As in the benchmark case, the savings choice is determined by trading off benefits and costs. Now, there is an additional term in the benefits reflecting the fact that the probability of being a buyer is endogenous and is captured by the cdf of money holdings. Thus, the entire distribution of money holdings matters and it arises as a solution of a differential equation.*

5 Remarks about fiat money and directed search

We have constructed a model of macroeconomic heterogeneity, inspired by Kiyotaki and Wright (1989) and driven by shocks to a linear-utility savings decision to hold a commonly desired medium of exchange. We have applied mechanism design with individual defections, and shown that in the benchmark case the optimum can be constructed by a static maximization problem. In that case, the optimum is not only time-invariant, but individual savings are actually invariant to the distribution of money.

Our *upper-bound argument* is so simple that it is tempting to conjecture that it may help demonstrating that the optimum in fiat-money models in the spirit of Lagos and Wright (2005) is also stationary. We concede that a fixed rate of money growth would translate into a fixed s in our model, and thus optimality could be stated relative to a fixed s. A full discussion of optimality, however, would have to address whether levels of s sufficiently close to β are feasible, like in the Friedman rule, and it is not clear that this is an interesting application of the upper-bound argument, at least because the optimum with near zero-opportunity cost of holding money is well known.[5]

Although stationarity is a natural prediction of linear models, and that is why our upper-bound argument is so fitting, the property that savings are invariant to distributions is not robust to a simple and appealing change in our model. In situations where a choice over consumer/producer status is physically feasible, allocating consumption to wealthy individuals is in fact socially optimal. We have shown that a differential equation can be used to describe the optimum in a class of smooth allocations. The shape of the optimal distribution of money reflects two forces. First, different people face different savings opportunities. Since output in meetings is divisible, there is an intensive margin to be explored, and a consequent dispersion in money holdings. This is the benchmark explanation for heterogeneity. Sec-

[5] The analysis of Lagos and Rocheteau (2008) also implies that the coexistence commodity and fiat monies is not interesting in linear models like ours. Commodity money would be driven out of the economy if the opportunity cost of fiat money can be made sufficiently low.

ond, because consumption opportunities are scarce, it is optimal to allocate consumption according to wealth, so that savers explore the distribution of holdings when making decisions. This is the auctions or extensive-margin explanation for heterogeneity.

We have thus linked heterogeneity and optimality. We recognize however that optimality need not be the only explanation for heterogeneity (and if frequently the least adopted in macroeconomics). An alternative can be found in the 'directed search' literature that builds on Burdett and Judd (1983). Their work focus on price dispersion resulting from precisely a lack of coordination by individuals that can sample prices from subsets of sellers. Julien et al. (2006) allow individuals to select sellers in a version of the Kiyotaki-Wright model with indivisible money, and consider an auction process to select the consumer according to bids in output quantities. This indivisibility restriction has been removed to allow bids of money by Galenianos and Kircher (2006), who also appeal to linearity in savings costs, but restrict consumption in meetings to be constant.

Two noticeable features of these monetary applications, in comparison to Burdett and Judd (1983), is that a seller can only serve one buyer at a time and, in equilibrium, buyers are allocated to sellers randomly. An econometrician observing pairs of sellers and (auction winner) buyers in their models, or pairs of auction traders in our model, could not tell which model is which. (If actual trades are observed, then the planner could 'disguise' our model by choosing, if necessary, implementable allocations with constant output or constant money transfers across meetings). Galenianos and Kircher (2006) provide nevertheless a promising avenue for future research, that could incorporate elements of optimality from our analysis. They offer an explicit discussion of asymmetric information in meetings (and a valuable literature review), which we have abstracted from completely. We conjecture that the upper-bound argument can be generalized in the presence of asymmetric information and a core-deviation concept.

References

[1] C. D. Aliprantis, G. Camera and D. Puzzello, Contagion equilibria in a monetary model, *Econometrica* **75** (2007), 277-282.

[2] F. Brauer and J. Nohel, *Introduction to differential equations with applications*, Harper & Row Publishers Inc., New York, 1986.

[3] K. Burdett and K. L. Judd, Equilibrium price dispersion, *Econometrica* **51** (1983), 955-969.

[4] R. Cavalcanti and A. Erosa, A theory of capital gains taxation and business turnover, *Economic Theory* **32** (2007), 477-496.

[5] R. Cavalcanti and P. Monteiro, On efficient distributions of money, mimeo.

[6] M. Galenianos and P. Kircher, A model of money with multilateral matching, *Journal of Monetary Economics* **55** (2008), 1054-1066.

[7] B. Julien, J. Kennes and I. King, Bidding for money, *Journal of Economic Theory* **142** (2008), 196-217.

[8] N. Kiyotaki and R. Wright, On money as a medium of exchange, *Journal of Political Economy*. **97** (1989), 927-954.

[9] R. Lagos and G. Rocheteau, Money and capital as competing media of exchange, *Journal of Economic Theory* **142** (2008), 247-258.

[10] R. Lagos and R. Wright, A unified framework for monetary theory and policy analysis, *Journal of Political Economy* **113** (2005), 463-484.

[11] S. Shi, A divisible search model of fiat money, *Econometrica* **65** (1997), 75-102.