Faculdade de Economia,
Administração e Contabilidade
de Ribeirão Preto
Universidade de São Paulo

Texto para Discussão

Série Economia

TD-E 14 / 2011

High interest rates: the golden rule for bank stability in Diamond-Dybvig model

Prof. Ms. Jefferson Donizeti Pereira Bertolai

Av. Bandeirantes, 3900 - Monte Alegre - CEP: 14040-900 - Ribeirão Preto - SP
Fone (16) 3602-4331/Fax (16) 3602-3884 - e-mail: cebelima@usp.br site: www.fearp.usp.br

Universidade de São Paulo

Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto
Avenida dos Bandeirantes,3900
14049-905 Ribeirão Preto – SP

A série TEXTO PARA DISCUSSÃO tem como objetivo divulgar: i) resultados de trabalhos em desenvolvimento na FEA-RP/USP; ii) trabalhos de pesquisadores de outras instituições considerados de relevância dadas as linhas de pesquisa da instituição. Veja o site da Comissão de Pesquisa em **www.cpq.fearp.usp.br**. Informações: e-mail: cpq@fearp.usp.br

High interest rates: the golden rule for bank stability in the Diamond-Dybvig model*

J. D. P. Bertolai and R. de O. Cavalcanti
FGV/EPGE

November 5, 2011

Abstract

In a companion paper, Bertolai et al. (2011) build on Peck-Shell (2003) economies and obtain strong implementation in perturbations of optimal contracts. Since bank runs are eliminated with distortions that become very small when the population grows, a pressing issue is whether an alternative specification can generate the costly crisis that are common in history. We find, in this paper, an affirmative answer in the context of the Diamond-Dybvig (1983) model, and uncover the role played by societal weights on future consumption and solvency risk. An extension of the Ennis-Keister (2009) algorithm shows the impact of run strategies and implicit rates of interest on the formation of expectations, in line with some classical views.
Keywords: severe aggregate-uncertainty, mixed-strategy bank runs, insolvency, dynamic programming. JEL codes: E4, E5.

1 Introduction

Perhaps less forgiving than modern counterparts, financial arrangements based on paper credit and fractional reserves, in the era of gold standard, gave to classical economists Adams Smith and Henry Thornton a special perspective on what today are called financial regulation and central banking. For Thornton, an illiquid system can accomplish more — despite the inherent risk of suspension of convertibility — when the industry seeks long-term payoffs, in contrast to narrow objectives associated to bank panics, high demand for gold and extinction of debt. One possible interpretation of Thornton ideas is that society should commit to policies making future growth attractive to a substantial set of agents, and that in doing so the potential for panics is minimized. There is, however, a considerable degree of ambiguity in the modern literature about the trade-offs involved in the provision of liquidity from a macro perspective. The purpose of this paper is to make a simple and transparent connection between bank stability and the social discounting of future utilities using variations of

*Preliminary and incomplete draft prepared for the Wisconsin School of Business, 3rd Annual Conference on Money, Banking and Asset Markets, November 2011.

the Diamond-Dybvig (1983) model. In particular, we discuss which implicit interest rates society should adopt when aggregate uncertainty is severe or when insolvency becomes a real threat.[1]

This paper builds on a literature inspired by Diamond and Dybvig (1983) but focusing on frictions in the diffusion of information that are fundamental. Wallace (1988) was the first to notice that a social planner should not attempt to treat *impatient* consumers equally in a banking model with sequential service and aggregate uncertainty. For him [see also Wallace (1990)], this basic property has deep roots, and should be interpreted as a single-asset explanation of actual bank suspensions observed in history. Since then, new tools have been developed, and in this paper we shall advance them to bring to bear constraints needed to insure truth-telling and uniqueness of equilibrium, two issues not pursued at the time. We find that variations in consumption at the first date of Wallace's sequential-service model is best interpreted as part of a movement in interest rates meant to give future utilities a proper weight and which should be modified by incentives and stability considerations. We also think that ideas for bringing insolvency issues into focus, which we also formalize, can increase our understanding of historic suspensions. We find, for instance, that in general the planner should not seek equal-treatment of *patient* consumers either.

Wallace's paper falls short of characterizing optimal allocations if depositors of different types — with different liquidity needs — arrive at random times to the bank. Green and Lin (2000 and 2003) reach much further, including examples that can be computed in close form. In more general specifications with independent shocks, reserves need not be distorted in order to provide incentives for truth telling and, remarkably, there is no need to worry about bank stability either: they prove that the optimum is interior and uniquely implemented in their economies. Peck and Shell (2003) propose a new landscape, however, showing that runs can become pervasive with new truth-telling constraints resulting from a less informed depositor (see also Andolfatto et al., 2007, for more on the role of disclosure).

Peck and Shell (2003) and, more recently, Ennis and Keister (2009), manage to open up several directions for future work, with examples of runs, design of strong implementation in pure strategies (Peck-Shell) and an algorithm for optimal reserves (possibly susceptible to runs) for a particular case of correlated shocks and inactive truth-telling constraints (Ennis-Keister). In Bertolai et al. (2011), however, the issue of strong implementation becomes very subtle when the distribution of types is easily predicted. They show that if the population is not too small and types are independent, then optimal reserves are essentially invariant to typical changes in truth-telling constraints (differences are particularly small in numerical examples with homothetic preferences).[2] This suggests

[1] In contrast to the proposals of matching maturity of assets and liabilities, by Adams Smith, or prohibition of fractional reserves, by Milton Friedman [see Wallace (1988) for references], Henry Thornton ([1802] 1939, chap. 4) warns against reductions in the creation of liquid liabilities by the central bank for preventing financial crises. An alternative, a policy of extending liquidity in exchange for good collateral, is largely credited to Walter Bagehot (1873), whose proposal includes a recommendation for higher "penalty" rates.

[2] A different but somewhat related subtlety is found by Cavalcanti and Monteiro (2011). They propose circumventing the revelation principle in order to achieve strong implementation in Ennis-Keister and Peck-Shell environments through large message-spaces, allowing for partial withdrawals and two-price mechanisms. Since, to some degree, such mechanisms are subject to the criticism of being unrealistic, we do not pursue them here.

that reserve management is more meaningful — eliminating runs is more costly — if aggregate uncertainty remains substantial even after the intentions of an initial subset of depositors can be sampled. In this paper, we shall provide an algorithm for strong implementation generalizing the weak-implementation method in Ennis and Keister (2009) in other respects: it allows for active truth-telling constraints in the spirit of Peck and Shell (2003), and for a different notion of aggregate uncertainty that avoids the limit result in Bertolai et al. (2011).

Our extension to insolvency risk in the Diamond and Dybvig (1983) model is also novel. The need to avoid a bank run becomes more appealing when financial distress can leave a number of depositors without consumption (or with very low utility). We borrow from Kocherlakota and Wallace (1998), and Peck and Shell (2010), monitoring assumptions that can lead to insolvency. In the former, a planner learns about actions in a monetary setting only after a time lag.[3] In the latter, a bank cannot store information provided by those willing to visit without withdrawing. In our formulation, patient individuals — with high desire for late consumption — are further divided into two groups: in one case (type 1), actions are monitored as in the conventional model, and in the other (type 2), a given individual is able to mimic the actions of a subset (of size two, in our numerical examples) of impatient depositors. Embezzlement behavior by a type-2 person at the first period, if it happens to occur in a run equilibrium, can be detected only in the second period, thus leaving the bank insolvent. We find that reserve management should direct extra second-period resources to these potential 'insiders' in order to promote bank stability.

Golden rule: less discounting, more savings

Our basic construction, employed throughout the paper, can be explained in simple terms for the particular case of independent shocks and standard monitoring (no type-2 people). There are two dates and two types of depositors. The aggregate state is $\omega \in \{0,1\}^N$, where N is the population size. A null coordinate i, $\omega_i = 0$, means that an individual arriving to the bank in position i consumes only at the first date (c_1) and has utility $Au(c_1)$, where $A \geq 1$ and $u(c)$ stands for $\frac{1}{1-\delta}c^{1-\delta}$ with $\delta > 1$. Otherwise, when $\omega_i = 1$, 'person i' can enjoy consumption at both dates (c_1 and c_2) and her utility is instead $u(c_1 + c_2)$. Individuals are ex-ante identical and later experience idiosyncratic shocks to preferences and to positions. The probability of drawing type 0 is p and, for any given type realization, the probability of position i is $\frac{1}{N}$. As in Peck and Shell (2003), individuals must announce types without knowing their positions; the planner — the bank — must transfer a quantity of date-1 consumption to the person in position i based on history $(\omega_1, ..., \omega_i)$ and taking expectations about $\omega_{i+1}, ..., \omega_N$ (sequential service). For a given i, date-2 consumption is a function of the whole list ω, and the aggregate date-2 consumption is bounded above by reserves saved at date 1 and reinvested at (gross) rate of return $R > 1$. The bank starts with initial reserves Y. A reserve-management rule should be designed so as to maximize ex-ante utility.

[3]See also Calomiris and Khan (1991), and Prescott and Weinberg (2003), as well as other references on the treatment of imperfect monitoring in monetary theory reviewed by Cavalcanti (2011).

Because $R > 1$, a necessary condition for welfare maximization is that type-1 individuals consume zero at the first date, enabling society to make the best use of its growth potential. Also, it can be shown that in this simple economy all type-1 individuals should share the same level of consumption at the second date (this also holds with binding constraints here but not in our general formulation of aggregate uncertainty below).

Ennis and Keister (2009) establish that optimal reserves for this homothetic economy follow from a dynamic-programming solution, exploring the fact that the marginal benefit of a unit saved to date 2 and shared with any j patient individuals can be found analytically. Also, the level of date-1 consumption that should be transferred to an impatient person at position i (defining reserves saved for position $i + 1$) is an explicit fraction

$$\frac{1}{1 + f_i(\omega_1, ..., \omega_i)}$$

of previous savings, where the coefficient f_i vary with the number of zeros in history $(\omega_1, ..., \omega_i)$. And, more importantly, Ennis and Keister (2009) show that these coefficients can be computed recursively.

One problem with the Ennis-Keister calculation is that it may discount future utility too much if A is sufficiently high. This is the case when types are private information, a situation assumed throughout this paper, and when A is sufficiently high. Increases in A lead to higher marginal utility of date-1 consumption and, if no corrections are made, a patient individual will choose to misrepresent her type in order to reach a better payoff. We show, in this paper, that the problem can be corrected as follows. Instead of assuming that the type-1 utility is $u(c_1 + c_2)$, the planner can apply the Ennis-Keister algorithm to an alternative economy where the patient utility function is now $\beta u(c_1 + c_2)$, and where β is a correction factor to be determined. With this abstraction in mind, we construct an algorithm that computes the slackness of truth-telling constraints associated to the Ennis-Keister program for the 'β-economy'. If for $\beta = 1$ the constraint slacks then the optimum has been found. Otherwise, a search for β is constructed, with increments that depend on the degree of individual gain that a type misrepresentation generates.

The previous exposition refers to weak implementation. Let us suppose now that a sought-after value β, organizing bank reserves according to the modified Ennis-Keister algorithm, has been found. By construction, a patient individual reveals her type under the assumption that others report truthfully. But nowhere in this design it has been taken into account what individuals prefer to do in the scenario of a bank run, when other patient depositors choose to withdraw. This leads to the question of how reserves should be designed within the class of mechanisms without multiple equilibria. Our approach is to incorporate into our method *no-run constraints* introduced by Peck and Shell (2003). Our version is broader, eliminating all runs in mixed, symmetric strategies. We do find another dynamic programming leading to the desired solution. It applies new shadow prices: date-1 consumption is socially valued according to $\alpha A u(c_1)$, while that for date 2 is valued according to $\beta u(c_2)$, for suitably chosen α and β that vary along the history but are still computable by a single-dimensional search protocol as before.

This intuitive approach is shown to be applicable to new questions alluded to above. The main conclusion is that sufficiently high (implicit) interest rates guarantee incentive compatibility and bank stability. But even within the basic environment of independent shocks, new findings are demonstrated. We show, for instance, that bank runs are pervasive. While Peck and Shell (2003) appeal to a strong demand for liquidity — setting the constant A equal to ten — in order to find bank fragility, we show that an increase in the supply of liquidity plays the same role, without additional assumptions. All that is needed is an increase in the population size, making insurance easier to provide. In particular, we find two equilibria besides the optimal one. Namely a run in pure strategies and a run in mixed strategies.

The distortions that make strong implementation possible, not surprisingly, are found to vary with insurance levels. We use simulations to document properties not explored in the literature to date. For instance, and again in the basic economy with independent shocks, the size of the economy produces two opposing effects. First, it facilitates the provision of insurance as a larger population can more easily pool risks. And, with higher insurance, run strategies become more attractive. Second, in order to discourage runs, an economy with many traders can spread distortions across many events. Hence, a distortion placed after a long sequence of withdrawals produces a small tax on the average depositor (a kind of 'backloading' property common in the public finance literature). The final effect confirms a seemingly paradoxical outcome: with independent shocks, runs can be removed in very large economies without substantial welfare losses for the average person. In particular, a run-proof management of reserves should tax heavily only the impatient arriving late to unusually long withdrawal events. This conclusion, however, is not valid when correlation is introduced, and we present a careful discussion for this reversal below.

The rest of the paper is divided as follows. Section 2 presents the basic environment with the standard monitoring assumption. Section 3 introduces a shadow-price approach to take distortions into account in simulations. Section 4 illustrates how the method can be used to characterize financial fragility. Section 5 extends the analysis to the design of run-proof reserves. Section 6 presents the environment with imperfect monitoring and the corresponding findings. Section 7 concludes.

2 The environment without insolvency

The benchmark economy in our analysis is hit by a shock ω with support $\Omega \equiv \{0,1\}^N$. The parameter N stands for the number of ex ante identical depositors that live for two dates and derive utility from pairs (c_1, c_2) of consumption provided by a bank—the benevolent social planner, who controls the aggregate endowment Y—according to positions and announcements about preferences that are private information. While for now an individual can be of two types, 0 or 1, in the last section of the paper we add a third type in order to discuss insolvency.

Person i is called *impatient* if $\omega_i = 0$ and called *patient* otherwise. The utility in the former case is $Au(c_1)$ and in the latter is $u(c_1 + c_2)$, where $A \geq 1$ and $u(c) = \frac{1}{1-\delta}c^{1-\delta}$ and $\delta > 1$. Thus only patient individuals can substitute consumption across dates. The resources not consumed in date 1 are reinvested at gross rate-of-return $R > 1$. These preferences have been used by Green and Lin (2000), with $A = 1$, and Peck and Shell (2003), with $A = 10$.

Feasible transfers must be incentive-compatible and satisfy a sequential-service constraint. The sequential-service constraint prevents date-1 consumption transferred to a person in position i to depend on information provided by someone at position n for $n > i$.

In the standard model, each individual draws a unique position i in $\{1, ..., N\}$ with probability $\frac{1}{N}$ and, as a result, the realization ω_i, without knowing the other coordinates of ω. As in Peck and Shell (2003), we shall assume that the individual is not informed of his position i. But we also want to consider a case in which shocks are correlated and let the degree of correlation vary in a simple manner, having independence as a particular case. The details about this more general stochastic process for types are given at the end of this section. In comparative-statics exercises with variations in population sizes, we shall keep the per capita endowment $e = \frac{Y}{N}$ constant as the population size, N, varies.

A compact description of candidates for optimal allocations follows from additional notation. Let us denote by ω^i the vector $(\omega_1, \omega_2, ..., \omega_i)$, and by (ω_{-i}, z) the profile that results from substituting the i-th coordinate of ω by z. Given that $R > 1$ we can restrict attention to transfers that assigns, to someone at position i, $x_i\left(\omega^{i-1}, 0\right)$ units of date-1 consumption, if that person is impatient, or $y_i(\omega)$ units of date-2 consumption, otherwise. The sequential-service requirement has thus shaped the domains of x_i and y_i. We shall denote by (x, y) a typical pair of transfer functions. We notice that (x, y) is feasible if

$$\sum_{i=1}^{N}\left((1-\omega_i)\, x_i\left(\omega^{i-1}, 0\right) + \omega_i R^{-1} y_i(\omega)\right) \leq Y, \tag{1}$$

and incentive-compatible if

$$E\left[\frac{1}{N}\sum_{i=1}^{N} u\left(y_i\left(\omega_{-i}, 1\right)\right)\right] \geq E\left[\frac{1}{N}\sum_{i=1}^{N} u\left(x_i\left(\omega^{i-1}, 0\right)\right)\right], \tag{2}$$

that is, when patient individuals that are not informed of their positions agree with revelation. [4]

The planner's problem is that of maximizing the representative-agent utility, before types and positions are assigned,

$$E\left[\frac{1}{N}\sum_{i=1}^{N}\left((1-\omega_i)\, Au\left(x_i\left(\omega^{i-1}, 0\right)\right) + \omega_i u\left(y_i\left(\omega\right)\right)\right)\right], \tag{3}$$

subject to (1) and (2).

A realization ω is interpreted as the type-composition of a queue for being served by the bank. We let the probability distribution on types for someone in position i depend on the realized type for the person in position $i-1$. In particular, the first-position person is impatient with probability p. And the person in position i is assumed to have the same type of the previous individual in the queue with probability θ. Accordingly, the probability distribution for position i on $\{0, 1\}$ is $(1-\omega_{i-1}, \omega_{i-1})$ times

$$T \equiv \begin{bmatrix} \theta & 1-\theta \\ 1-\theta & \theta \end{bmatrix}.$$

[4] The expectations in (2) are taken with respect to the distribution of ω_{-i} on $\{0,1\}^{N-1}$, given that $\omega_i = 1$.

Therefore, a history ω has probability

$$P(\omega)=\bar{\omega}_1\left[\begin{array}{c}p\\1-p\end{array}\right]\prod_{i=2}^{N}\bar{\omega}_{i-1}T\bar{\omega}_i'^{1-\omega_1}(1-p)^{\omega_1}\left[\theta^{N-1-s(\omega)}(1-\theta)^{s(\omega)}\right]$$

if $s(\omega)=\sum_i^{N-1}|\omega_{i+1}-\omega_i|$ is the number of type switches and $\bar{\omega}_i=\left(\begin{array}{cc}1-\omega_i & \omega_i\end{array}\right)$.

3 The Lagrangian approach

The planner's problem defined above aims to find the best incentive-compatible mechanism (x,y). As it will become clear further below, every (x,y) defines a game of announcements and it can happen that this game have multiple equilibria, with the intended revelation-equilibrium being just one of them. But before we can address multiplicity, it is important to be able to compute the optimal (x,y) defined above. While Green and Lin (2000) finds a closed form solution for $N=3$ and independent shocks, Ennis and Keister (2009), henceforth EK, derive a recursive method that handles correlated shocks (although with a different specification compared to the setup above). Our objective is to extend the Ennis-Keister (2009) approach. But one problem is that, like in the situation addressed by Green and Lin (2000), the incentive constraints could be ignored, while this is not the case for (2) above when A and θ are arbitrary. We shall see, however, that an intuitive extension is possible if we shift the attention away from the objective (3) and focus instead on a lagrangian version.

3.1 The objective for weak implementation

Proposition 1 *The optimum can be computed by [ignoring (2) and] replacing the objective (3) with*

$$\max_{(x,y)}\sum_{\omega}P(\omega)\sum_i\left[(1-\omega_i)\left(\alpha_i^{\omega_{i-1}}\right)^{\delta}u(x_i)+\omega_i\left(\beta_i\right)^{\delta}u(y_i)\right],$$

where $(\alpha_i,\beta_i)_{i=1}^N$ are functions of ω [determined in close form for each candidate Lagrange multiplier for (2)].

Proof. A patient agent when in position i believes that announcement profile ω_{-i} happens with probability

$$\Pr(\omega_{-i}|\omega_i=1)=\frac{\Pr(\omega_{-i},[\omega_i=1])}{\Pr([\omega_i=1])}=\frac{P(\omega_{-i},1)}{\sum_w P(w_{-i},1)}=\frac{P(\omega_{-i},1)}{\sum_{w^{i-1}}P(w^{i-1},1)}$$

where the denominator can be written as $q_i=(p,1-p)T^{i-1}\left([0,1]'\right)^5$. Therefore, the incentive constraint (2) can be rewritten as

$$\frac{1}{N}\sum_i\frac{1}{q_i}\sum_{[\omega:\omega_i=1]}P(\omega)\Big[u\left(y_i\left(\omega_{-i},1\right)\right)-u\left(x_i\left(\omega^{i-1},0\right)\right)\Big]=$$

$$\frac{1}{N}\sum_{\omega}P(\omega)\sum_i\frac{1}{q_i}\Big[\omega_i u\left(y_i\left(\omega_{-i},1\right)\right)-(1-\omega_i)\gamma_i^{\omega_{i-1}}u\left(x_i\left(\omega^{i-1},0\right)\right)\Big]\geq 0$$

[5]Note that $p=0.5$ implies that $q_i=0.5$ for all i

where, for $t \in \{0,1\}$, γ_i^t stands for $((1-\theta)/\theta)^{1-2t}$ when $i > 1$, and for $(1-p)/p$, otherwise. The term $\gamma_i^{\omega_{i-1}}$ is actually equal to the ratio between $P(\omega^{i-1},1)$ and $P(\omega^{i-1},0)$ which appears in the expression above when the summation is taking place over ω in Ω [as in (3)] instead of ω in $\omega \in \Omega : \omega_i = 1\}$ [as in (2)]. After this adjustment the slack in the truth-telling constraint, which appears in the lagrangian, becomes written as a summation with weights $P(\omega)$ as in the objective function.

If λ denotes the lagrangian multiplier for (2) then the planner's problem can be stated as

$$\max_{(x,y)} \left\{ \frac{1}{N} \sum_{\omega} P(\omega) \left(\sum_i \left[(1-\omega_i) \left(\alpha_i^{\omega_{i-1}}\right)^{\delta} u(x_i) + \omega_i \left(\beta_i\right)^{\delta} u(y_i) \right] \right) ; (1) \right\}$$

where $\alpha_i^t = \left(A - \frac{\lambda}{q_i}\gamma_i^t\right)^{1/\delta}$ and $\beta_i = \left(1 + \frac{\lambda}{q_i}\right)^{1/\delta}$. Notice that the q's and γ's can be computed using only properties of the distribution of partial histories, and nothing else (they are invariant to λ). Hence the coefficients α's and β's are computed in close form, and the proof is now complete. ■

3.2 Candidate policy functions

Now the EK dynamic-programming approach can be applied to solve the modified problem for each candidate multiplier. In effect, consider the date-2 partial problem faced by the planner after history ω:

$$\max_{y} \left\{ \sum_i \left(\beta_i\right)^{\delta} u(y_i); \sum_i \omega_i y_i(\omega) \leq Ra \right\},$$

where a denotes the sum of resources not consumed in date 1. Its solution must satisfy $y_i = \beta_i/\mu^{1/\delta}$ if μ is the multiplier for the resource constraint. Since it binds at the optimal solution, then $\mu^{1/\delta} = (\sum_i \omega_i \beta_i)/Ra$, which yields

$$y_i = \frac{\beta_i}{\sum_k \omega_k \beta_k} Ra. \tag{4}$$

Therefore, the corresponding optimal value is $(f_N(\omega))^{\delta} u(a)$, where $f_N(\omega) = R^{1/\delta - 1} \sum_k \omega_k \beta_k$. Keeping this contingent-solution fixed, the planner faces after position $N-1$ the partial problem

$$\max_{c \leq a} \left\{ T_{t+1,1} \Big(\left(\alpha_N^t\right)^{\delta} u(c) + f_N(\cdot, 0)^{\delta} u(a-c) \Big) + T_{t+1,2} \Big(f_N(\cdot, 1)^{\delta} u(a) \Big) \right\}$$

where a now denotes resources not consumed at the first $N-1$ positions, $t = \omega_{N-1}$, and $T_{i,j}$ denotes the element (i,j) in matrix T. The solution is now

$$c = \frac{\alpha_N^t}{\alpha_N^t + f_N(\cdot, 0)} a,$$

which produces the corresponding value

$$u(a) \Big(T_{t+1,1} [\alpha_N^t + f_N(\cdot, 0)]^{\delta} + T_{t+1,2} [f_N(, \cdot, 1)]^{\delta} \Big) = u(a) \Big(f_{N-1}(\cdot) \Big)^{\delta}.$$

Similar algebra can be done to show that the optimal solution for $i < N-1$ is always

$$c_i = \frac{\alpha_i^t}{\alpha_i^t + f_i(\cdot, 0)} a$$

and, likewise, the optimal value is $u(a)\Big(f_{i-1}(\cdot)\Big)^\delta$. The f's are coefficients for 'splitting the pie' and are functions of partial histories. They are moreover fully determined by the system given by

$$f_i(\cdot) = \Big(T_{t+1,1}[\alpha_{i+1}^t + f_{i+1}(\cdot, 0)]^\delta + T_{t+1,2}[f_{i+1}(\cdot, 1)]^\delta\Big)^{1/\delta}, \tag{5}$$

if $i > 0$, and

$$f_0 = \Big(p[\alpha_1 + f_1(0)]^\delta + (1-p)[f_1(1)]^\delta\Big)^{1/\delta}, \tag{6}$$

otherwise. The value for objective in the lagrangian/planner problem is therefore $N^{-1}u(Y)[f_0]^\delta$.

3.3 Iterating on multipliers

Unlike EK, we must verify whether a given discount β is correct. This is done by computing (2) under the implied allocation. If the incentive constraint is violated (slacks), β must be increased (reduced). The following proposition establishes a recursive method for computing the slack in the constraint. In the same way that we have found policies $f_i(\omega^i)$, determining consumption as a fraction of existing reserves at partial history ω^i, we shall look for partial sums in the expectations that define the slack in truth-telling constraints, and organize these terms by partial histories, using the notation $g_i^{\omega^i}$ as follows. Let us inspect the expression for the slack for given ω, as in the proof of Proposition 1, isolate the term

$$\sum_i \frac{1}{q_i}\Big[\omega_i u\left(y_i\left(\omega_{-i}, 1\right)\right) - (1-\omega_i)\gamma_i^{\omega_{i-1}} u\left(x_i\left(\omega^{i-1}, 0\right)\right)\Big]$$

and, in particular, consider the part associated to date-2 consumption. We write

$$g_N^\omega = \sum_i \frac{1}{q_i}\omega_i u\left(y_i\right)$$

where y_i is given by (4) for ω.

Proposition 2 *For each history ω, define $g_N^\omega = \sum_i \frac{\omega_i}{q_i}\left(\frac{\langle\omega,\beta\rangle}{\beta_i R}\right)^{1/\delta}$. If $g_i^{\omega^i}$ is defined as*

$$T_{t+1,1}\left[g_{i+1}^{(\cdot,0)}\left(1 + \frac{\alpha_{i+1}^t}{f_{i+1}(\cdot, 0)}\right)^{\delta-1}\right] + T_{t+1,2}\left[g_{i+1}^{(\cdot,1)} - \frac{1}{q_{i+1}}\left(1 + \frac{f_{i+1}(\cdot, 0)}{\alpha_{i+1}^t}\right)^{\delta-1}\right]$$

for $i > 0$, and

$$p\left[g_1^{(0)}\left(1 + \frac{\alpha_1}{f_1(0)}\right)^{\delta-1}\right] + (1-p)\left[g_1^{(1)} - \frac{1}{q_1}\left(1 + \frac{f_1(0)}{\alpha_1}\right)^{\delta-1}\right]$$

for $i = 0$, then $N^{-1}u(Y)g_0$ equals the slack in the truth-telling constraint when evaluated at the optimal mechanism implied by β.

Proof. The patient's payoff, in terms of an average across positions, given ω, is $u(a)g_N^\omega$ divided by N, where a is the endowment saved for date 2. Using now (4), we have $g_N^\omega = \sum_i \frac{\omega_i}{q_i}\left(\frac{\omega\beta}{\beta_i R}\right)^{\delta-1}$. A patient person can now proceed to compute the slack in his or her constraint according to deviation payoffs which vary across positions. We can compute the slack by restating the payoff $u(a)g_N^\omega$, with adjustments for either the opportunity from lying or for the law of motion of reserves, and according to a series of contingencies. If this person has drawn position N ($\omega_N = 1$) the corresponding expression is

$$g_N^{(\cdot,1)}u(a) - \frac{1}{q_N}u(x_N)$$

where x_N is the date-1 transfers at $(\omega^{N-1}, 0)$. If this person has not drawn position N ($\omega_N = 0$) , there is still an impact of x_N on available reserves that, when taken into account yields

$$g_N^{(\cdot,0)}u(a - x_N).$$

Integrating over these two contingencies using the Markov chain gives us a function of ω_{N-1} compactly written as

$$T_{t+1,1}\Big(g_N^{(\cdot,0)}u(a - x_N)\Big) + T_{t+1,2}\Big(g_N^{(\cdot,1)}u(a) - (1/q_N)u(x_N)\Big)$$

where a is now the endowment saved after the first $N-1$ positions and $t = \omega_{N-1}$. Now, using optimal value for x_N this expression amounts to $u(a)g_{N-1}^{\omega^{N-1}}$ as in the statement of the proposition.

Similar and straightforward algebra can be used to show that this partial measure of incentives to tell the truth for position $i < N$ equals $u(a)g_{i-1}^{\omega^{i-1}}$ if a is the quantity saved for position i, as in the statement of the proposition. ■

The proof demonstrates that one can start with the measure g_N^ω , summarizing the expected payoff from telling the truth at ω, and then recursively compute adjustments according to the what happens if the patient person draws each potential position. Integrating over all contingencies is facilitated by the Markov assumption. In summary, date-1 consumption affects the incentives of a patient person by reducing resources saved for date 2 and by its effect on misrepresentation payoffs. The former is accounted by the factor multiplying $g_{i+1}^{(\cdot,0)}$ in the first term of $g_i^{\omega^i}$ in the statement of the proposition, while the latter is the expression being subtracted in the second term.

The proposition gives us g_0 as a proxy for the slack on truth-telling constraints, which can then guide increases in implicit interest rates through changes in the multiplier λ (proposition 1). The EK method is, in this way, extended along the lines of penalty-function methods. The algorithm performs a simple search, by bracketing the positive orthant, until either $g_0 = 0$ is found or $g_0 > 0$ for $\lambda = 0$.

4 Financial stability

In this section we study the existence of run equilibria, and the cost in avoiding them. The optimal mechanism defines a game of announcements where patient agents can lie with probability π. *Run* equilibria are symmetric Bayesian-Nash equilibria of this game such that $\pi > 0$. In order to study the existence of runs for a fixed mechanism (x, y), we let for $\pi \in [0, 1]$

$$w(\pi) = \frac{1}{N} \sum_i E_\pi \Big[\omega_i u(y_i(\omega)) - (1 - \omega_i) u(x_i(\omega^{i-1}, 0)) \Big] \tag{7}$$

denote the relative payoff of truth-telling when other patient individuals are lying with probability π. In terms of the signal of the function w, a patient individual tells the truth (runs) if w is positive (negative), and the best reply to π is the interval $[0, 1]$ when $w(\pi) = 0$. Notice that w is a property of a weakly implementable (x, y) that describes its potential fragility, and that $w(0)$ can be computed according to the construction of g_0 derived above and the corresponding iterations of multipliers in our algorithm.

4.1 Pervasive runs

The computation of $w(\pi)$ for $\pi > 0$, as in Figure 1, is a by-product of the algorithm for strong implementation outlined below.

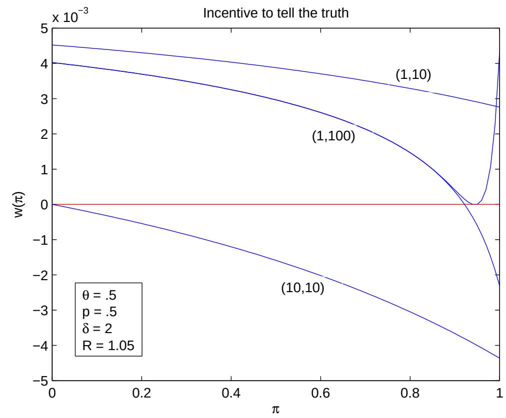

Figure 1: A mixed-strategy run equilibrium

Figure 1 presents the function w for some economies with *iid* shocks, which is the case when $\theta = .5$ and $p = .5$ (we keep $p = .5$ throughout the paper). The numbers in parenthesis correspond to the values of the pair of parameters (A, N). The three decreasing curves are representations of weak implementation discussed so far. The other curve, which bounces back after a tangency at the horizontal axis, has to be explained later because it corresponds to a w after judiciously selected distortions on (x, y) are introduced and which accomplishes strong implementation.

In these examples, weak implementation with independent shocks feature $w(\pi)$ decreasing in π because a patient person thinks that telling the truth becomes increasing unattractive as more people are running, leaving less and less resources left for date 2. The fact that reserves are planned to be constantly reduced with a sequence of zeros — say, along the contingency $\omega = (0, 0, ...)$ — is easily understood when we take into account the role of homothetic preferences.

After a given sequence of zeros, the planned transfers are computed as if future shocks are drawn from P, and not from a transformation of P generated by run strategies. Since preferences are homothetic, the drop in reserves up to this point has no effect on how the pie is planned to be divided from this point onwards. As result, a plan of providing insurance based on P for the remaining traders is maintained, which means that some sharing of the growth potential R with impatient people will continue to drive reserves down during a run. A key question is thus whether the function w eventually crosses into negative territory, when panic is widespread and π approaches one, demonstrating that an equilibrium-run exists. Intuitively, one factor leading to w negative is the level of liquidity insurance 'demanded' by the impatient. This factor has been emphasized by Peck and Shell (2003), with A increasing from 1 [as in Green and Lin (2000,2003)] to 10, and by EK, with increases in δ across examples. But the preceding discussion suggests that N also plays a big role. For, if N is large and shocks are independent, more insurance is planned and the consequent depletion of reserves lasts longer as the sequence of 0-announcements keep coming.

It can also happen that w is everywhere negative but at point $\pi = 0$. This case is illustrated by the lowest curve, which corresponds to a similar parametrization in Peck and Shell (2003, appendix B). The optimal allocation is constrained since $w(0) = 0$, and there is a run equilibrium in pure strategies since $w(1) < 0$. The highest curve shows that the the best-reply correspondence changes, and no runs are found, in the same economy, when $A = 1$. This exercise confirms that run equilibria can be obtained in small population economies by increasing A. And, in the another direction, one can increase the population size N to 100 to show that runs do not require a high value for A (neither a high value for δ used in the run example constructed by EK with independent shocks). The conclusion is that runs are always motivated by too much risk sharing. In addition, the large population example illustrates that there are mixed-strategy equilibria whenever the optimum is unconstrained and there is a run equilibrium in pure strategies.

The role of N in these examples take us to the limit result in Bertolai et al. (2011). They show that welfare of a Peck-Shell economy with independent shocks approaches that of a of an economy with a continuum of people as N grows to infinity. Their proof does not rely on homothetic preferences nor bounds on risk aversion. And moreover, they show the truth-telling constraint of the 'continuum' problem is active if and only if $A \geq R$. Bearing in mind this characterization of large-population economies, we guess and verify, numerically, the following claim for our homothetic economy, given the intuition constructed for the law of motion of reserves during a run.

Claim Assume that shocks are independent. There exist mixed-strategy runs in all economies with $A < R$ and N sufficiently high. On the converse, there are no mixed-strategy runs in economies with $A \geq R$.

4.2 Persistence and weak implementation

Because persistence introduces a load on memory requirements for numerical work that grows significantly with N, we have not attempted to study limit properties with $\theta = .75$ in general. The characterization of economies featuring active constraints is more subtle in this case, and the theorem in Bertolai et al. (2011) cannot serve as guidance. But we do find that the propensity to run increases when N is fixed and θ shifts from .5 to .75, for N is a large range. This tells us that mixed-strategy runs are still common, but the way that (A, R, N) determine the slack in truth-telling constraints changes.

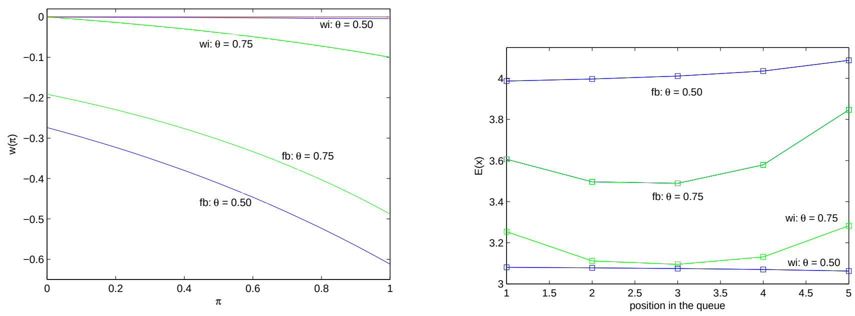

Figure 2: Aggregate-uncertainty effects.

Figure 2 illustrates the effects of introducing aggregate uncertainty with Markov transitions. The parameter values are the same as before, but $(A, N) = (10, 5)$. The two lowest curves in the graph on the left refers to first-best (fb) allocations for the cases of independence ($\theta = .5$) and persistence ($\theta = .75$). These are mechanisms obtained by ignoring truth-telling constraints, as if types are no longer private information. The other curves refers to weak implementation (wi).

Let us fix attention to the first-best curves on both graphs momentarily. For both θ's the first best is not implementable since $w(0) < 0$. In addition, the fb-w is lower (misrepresentation is more attractive) when there is less persistence because there is more room for risk sharing. Intuitively, there is more mixing in P. The graph on the right confirms such intuition: it can be seen that the average date-1 consumption in the first-best allocation is higher when shocks are independent. It can also be noticed that transfers are more sensitive to positions on average when shocks are persistent because current shocks are very important for predicting the future, even close to the end of the queue. That is, the sequential-information friction proposed by Wallace (1988) becomes stronger under correlation. A novel result, that we shall document below, is that Lagrange multipliers fall under correlation and, in the case of active constraints, there is again in the capacity of providing insurance because weak-implementaion distortions can be reduced. The reason for the fall in multipliers is explained below. We should see, however, that the corresponding increase in the provision of insurance makes runs more likely.

When forced to provide incentives to patient agents to tell the truth — when forced to change (x, y) so that $w(0) \geq 0$) — the planner distorts the allocation in a way that misrepresentation becomes more attractive for higher θ. The reason for this inversion in w's can be found in the second graph of figure 2: planner is able to provide incentives in the persistence case with a smaller reduction in date-1 consumption, so that the deviation payoff does not decrease as much as in the independent case. Intuitively, it is easier to convince patient individuals to not run when shocks are persistent. In effect, due to persistence, a patient person knows that with high probability announcements will be mostly patient ones. Telling the truth in this scenario is good since date-2 payments are high. On the other hand, the deviation payoff for this individual is mainly concentrated (in a probability sense) on withdrawals just after a patient announcement. Knowing this, the planner can concentrate distortions on impatient announcements which happens after a type innovation. That way, the planner is able to preserve a high insurance level on average since distortions placed on withdrawals immediately after impatient announcements can be made temporary, easily preventing lies (again in a probability sense).[6]

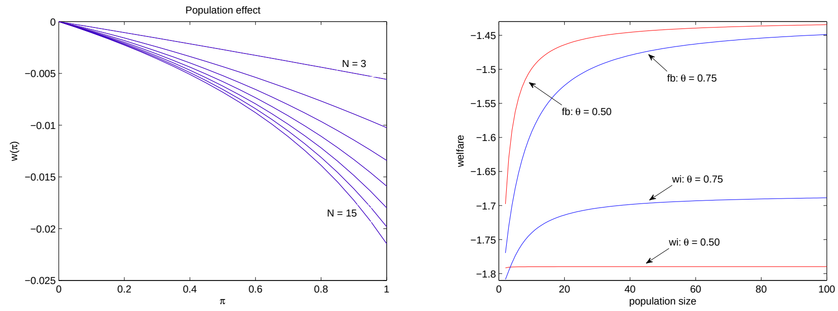

Figure 3: Population effects.

Figure 3 illustrates the effects of population growth on economies with persistence. The graph on the right shows that the ability to maintaining risk sharing under weak implementable allocations plays an important role when the population increases. Welfare significantly improves even in a constrained optimum when $\theta = .75$, but not so much when $\theta = .5$. If there was no incentive problem, insurance would increase in both cases as can be concluded from first-best curves (fb) in the second graph. When incentives are taken into account (wi), the curve for $\theta = .5$ confirms that the limit result in Bertolai et al. (2011) occurs for very low populations with Peck-Shell preferences: in the scale presented by the graph, welfare with small N coincides with that of a very large economy. The corresponding speed of convergence is much lower when $\theta = .75$. We shall see that, although welfare eventually flattens up with persistence, the economy with $\theta = .75$ is qualitatively different because the incentives to run in a panic are magnified, and the remedy cost is maintained considerable as the population grows, in contrast to the perturbation result in Bertolai et al. (2011). There, with independent shocks, the cost of removing panics converges to zero exponentially as the population grows.

When $\theta = .75$, increases in population is reflected in smaller and smaller values for w.[7] The graph on the left of Figure 3 shows such shifts in w, and

[6] More formally, we notice, from proposition 1, that date-1 consumption has weight $\alpha_i^t = \left(A - (\lambda/q_i)\gamma_i^t\right)^{1/\delta}$ in the modified problem that planner solves under weak implementation. Because in general $\gamma_i^t = (1/\theta - 1)^{1-2t}$, such weight is invariant to t when shocks are independent ($\theta = .5$), while it is high for $t = 0$ and low for $t = 1$ when there is persistence ($\theta > .5$).

[7] We have seen in Figure 1 that a similar effect is occurring when shocks are independent, which is tatonnement to the pervasiveness of run equilibria documented in the claim that follows the figure.

we have done the same exercise for $\theta = .5$ (not shown in the graph). In both cases, the shifts become smaller as N increases, but they are much larger when θ is high. In addition, w is always much smaller in the persistent case, with $w(1) \approx -10^{-2}$ for $N = 11$ in the independent case. As discussed above, the impact of increases in N on deviation payoffs is driven by the insurance effect. It is always higher in the persistence case because the planner can induce truth-telling at a lower cost, allowing for higher date-1 consumption on average. As the population grows, the distortions needed to provide incentives under weak implementation are reduced, and the consequent impact on the provision of insurance causes shifts on the w's curves that keep changing, even though in lower rates in both cases.

4.3 The cost of eliminating runs

Next we address the issue of what is the least costly way to promote distortions in (x, y) in order to eliminate runs. We propose to measure the 'welfare cost of financial fragility' by studying the effects of imposing 'no-run constraints' that have been defined in Peck and Shell (2003) for a context of pure strategies. We shall approach this matter by extending the concept to mixed strategies, and extending the numerical work to cover these constraints simultaneously with the standard truth-telling constraints. The new method is responsible for the construction of the non-monotone w curve in Figure 1 which remains always nonnegative.

We call *cost of financial stability* the reduction in the per capita endowment that has to applied in order to produce an optimal welfare, under weak implementation, of the same magnitude as that under strong implementation. In other words, we seek the least costly way to introduce distortions in the economy in order to eliminate runs. We have seen that deviation payoffs are higher when there is more persistence and when the population is larger. We recall that Bertolai et al. (2011) construct a simple mechanism featuring strong implementation and which converges to the weakly-implementable optimum when $N \to \infty$ and shocks are independent. Hence, we must confirm numerically, using the Peck-Shell concept of no-run constraints, that the cost of financial stability indeed converges to zero when $\theta = .5$.

The methods developed so far generates a w function such that $w(0)$ is nonnegative. The necessary changes are guided by the following definition.

Definition 3 *The optimal allocation associated to a revelation equilibrium ($\pi = 0$) for a given mechanism is strongly implementable if the w for this mechanism is nonnegative for all π.*

The idea is to force, if necessary, a better treatment of patient individuals so that, generically, runs are not best replies. In particular, we seek to solve the following.

SI problem Maximize (3) in (x, y), subject to (1) and to $w(\pi) \geq 0$ for all $\pi \in [0, 1]$.

The following results shows that a simple extension of the current algorithm is able to solve the SI problem.

Proposition 4 *Assume that the solution to the SI problem features $w(\pi) = 0$ at finitely many π's and that the respective multipliers λ_π's are known. Then the this solution can be computed by modifying the objective as*

$$\sum_\omega P(\omega) \sum_i \left[(1-\omega_i)\alpha_i(\omega^i)^\delta u(x_i) + \omega_i \beta_i(\omega)^\delta u(y_i)\right]$$

where

$$\alpha_i(\cdot) = \left(A - \frac{1}{q_i}\int_0^1 \lambda_\pi h_\pi^i(\cdot, i) d\pi\right)^{1/\delta}$$
$$\beta_i(\cdot) = \left(1 + \frac{1}{q_i}\int_0^1 \lambda_\pi h_\pi(\cdot, i) d\pi\right)^{1/\delta}$$

and $(h_\pi, h_\pi^i)_{i=1}^N$ are functions of ω determined in close form.

Proof. Let η stand for the history of announcements. A patient agent believes that an impatient announcement happens in position i with probability

$$T_{t+1,1}^\pi = T_{t+1,1} + \pi T_{t+1,2} = \langle \bar{\omega}_{i-1} T, (1, \pi)\rangle,$$

if $i > 1$, and $\langle (p, 1-p), (1, \pi)\rangle$, otherwise. Accordingly,

$$P_\pi(\eta) = \sum_\omega P_\pi(\eta|\omega)P(\omega) = \sum_\omega P(\omega)\prod_{i=1}^N \bar{\omega}_i M_\pi \bar{\eta}_i'$$

If in position i, he or she believes that announcement η_{-i} happens with probability

$$\begin{aligned}
\Pr(\eta_{-i}|\omega_i = 1; \pi) &= \frac{P_\pi(\eta_{-i}, [\omega_i = 1])}{\Pr([\omega_i = 1])} = q_i^{-1}\sum_\omega P_\pi(\eta_{-i}, [\omega_i = 1]|\omega)P(\omega) \\
&= q_i^{-1} \sum_{[\omega:\omega_i = 1]} P_\pi(\eta_{-i}|\omega)P(\omega) \\
&= q_i^{-1} \sum_{[\omega:\omega_i = 1]} P(\omega)\left[\prod_{j\neq i}\left(\bar{\omega}_j M_\pi \bar{\eta}_j'\right)\right] = \frac{P(\eta_{-i}, t)}{q_i} h_\pi((\eta_{-i}, t), i)
\end{aligned}$$

where $M_\pi = \begin{bmatrix} 1 & 0 \\ \pi & 1-\pi \end{bmatrix}$ and

$$h_\pi((\eta_{-i}, t), i) = \sum_{[\omega:\omega_i = 1]} \frac{P(\omega)}{P(\eta_{-i}, t)}\left[\prod_{j\neq i}\left(\bar{\omega}_j M_\pi \bar{\eta}_j'\right)\right]. \tag{8}$$

That way, we have

$$\begin{aligned}
w(\pi) &= \sum_i \frac{1}{N}\sum_{\eta_{-i}} \Pr(\eta_{-i}|\omega_i = 1; \pi)\left[u(y_i(\eta_{-i}, 1)) - u(x_i(\eta^{i-1}, 0))\right] \\
&= \frac{1}{N}\sum_i q_i^{-1}\sum_\eta P(\eta)h_\pi(\eta, i)[\eta_i u(y_i(\eta_{-i}, 1)) - (1-\eta_i)u(x_i(\eta^{i-1}, 0))] \\
&= \frac{1}{N} E \sum_i q_i^{-1}\Big[\omega_i h_\pi(\omega, i) u\left(y_i\left(\omega\right)\right) - (1-\omega_i)h_\pi^i(\omega^i, i)u\left(x_i\left(\omega^i\right)\right)\Big]
\end{aligned}$$

where expectation is now taken with respect to ω, and the h's are given by

$$h_\pi^k(\cdot,i) = \left\langle \bar{\omega}_k T, \left(h_\pi^{k+1}((\cdot,0),i), h_\pi^{k+1}((\cdot,1),i)\right)\right\rangle$$

for each ω^k if $h_\pi^N(\omega,i) = h_\pi(\omega,i)$.

If λ_π denotes the lagrangian multiplier on (7), then SI problem can be stated as solving

$$\max_{(x,y)} \left\{ \frac{1}{N} \sum_\omega P(\omega) \left(\sum_i \left[(1-\omega_i)\alpha_i(\omega^i)^\delta u(x_i) + \omega_i \beta_i(\omega)^\delta u(y_i) \right] \right); (1) \right\}$$

and the proof is now complete. ■

Again the EK dynamic-programming approach can be applied to solve such a problem for each candidate multiplier set, $\boldsymbol{\lambda} = \{\lambda_\pi : \pi \in [0,1]\}$. The optimal solution is still given by (5-6), but now α_i depends on ω^i, not only on ω_{i-1} as in the weak-implementation case. It turns out that $w(\pi)$ can be computed recursively in the same fashion as done for the slack in the truth-telling constraint in Proposition 2.

Proposition 5 *For a fixed belief π, define $g_i^{\omega^i}$ as*

$$T_{t+1,1}\left[g_{i+1}^{(\cdot,0)}\left(1+\frac{\alpha_{i+1}(\cdot,0)}{f_{i+1}(\cdot,0)}\right)^{\delta-1}\right] + T_{t+1,2}\left[g_{i+1}^{(\cdot,1)} - \frac{h_\pi^i((\cdot,1),i)}{q_{i+1}}\left(1+\frac{f_{i+1}(\cdot,0)}{\alpha_{i+1}(\cdot,0)}\right)^{\delta-1}\right]$$

for $0 < i < N$, and

$$p\left[g_1^{(0)}\left(1+\frac{\alpha_1}{f_1(0)}\right)^{\delta-1}\right] + (1-p)\left[g_1^{(1)} - \frac{h_\pi^1((1),1)}{q_1}\left(1+\frac{f_1(0)}{\alpha_1}\right)^{\delta-1}\right]$$

for $i=0$, together with $g_N^\omega = \sum_i \frac{\omega_i h_\pi(\cdot,i)}{q_i}\left(\frac{\langle\omega,\beta\rangle}{\beta_i R}\right)^{1/\delta}$. Then $N^{-1}u(Y)g_0$ equals $w(\pi)$ when evaluated at the solution associated to $\boldsymbol{\lambda}$.

Proof. The argument is identical to the one used in Proposition 2, after the adjustment required by the introduction of the terms h and h^i, and is thus omitted. ■

The non-monotone curve in Figure 1 illustrates the typical case found by the algorithm when mixed-strategy runs are rule out. In this case there is only on critical π to be found (and only one multiplier for the solution): it corresponds to the pivotal (in the language of linear programming) strategy for which the corresponding no-run constraint is weakly active. Once this critical π is found, all other constraints slack. And because the resulting non-monotone curve is positive at 0, the standard incentive constraint (implied by the revelation principle) also slacks. We conclude that the machinery used here to address mixed strategies are important for an accurate assessment of the costs of financial stability.

In order to quantify the distortion in the optimal allocation, we compute by how much the per capita endowment should be reduced under weak implementation to achieve the welfare under strong implementation. The following table summarizes the results for selected economies.

θ	Endowment reduction $N=3$	$N=5$	$N=7$	Interest rate: $E(y)/E(x)$ first-best	weak	strong
.50	.0022%	.0021%	.0017%	.4027	.9258	.9262
.75	.1543%	.3173%	.3786%	.3572	.5874	.6092

The left side of the table shows that for all considered population sizes, cost of financial stability is much larger when there is persistence. We also confirm that the cost is very low in the independent case. While in Bertolai et al. (2011) there was a noticeable convergence of Peck-Shell and Green-Lin mechanisms (when there is disclosure of positions and traders use elimination of dominated strategies, in computed cases, to choose truth-telling) for such low N's. Hence, because there were no runs for the Green-Lin setting in the numerical examples of Bertolai et al. (2011), the conclusion was that by letting traders be informed of their positions in the queue — if such disclosure can be considered feasible, what is debatable — the planner can achieve strong implementation at a low cost (since the transfer functions of Peck-Shell and Green-Lin mechanisms quickly converge to each other and there no runs with disclosure). Here, by contrast, we show that disclosure is not needed. By placing the appropriate distortions, according to our extended algorithm, we find that the cost falls to zero very quickly when $\theta = .5$. The effects of an increase in θ on stability costs are very strong because the distortions needed are more severe when there is a threat that a whole group of people is running (the case $\pi > 0$ that has to be addressed by strong implementation). With persistence, when forced to provide incentives to patient agents to not run, when they believe other patient individuals are misrepresenting (when forced to change (x, y) so that $w(\pi) \geq 0$ for all $\pi \in [0, 1]$), the planner must distort withdrawals in a larger range of 0's. Intuitively, a deviator thinks that he or she will be bunched with other zeros. Hence, a belief $\pi > 0$ shifts (in a probabilistic sense) the deviation payoffs compared to the weak-implementation case. Now there is a need to reassure patient people whose deviation payoff accrues after a type innovation, as well as those whose deviation payoff accrues after a type persistence. Thus the planner must now reduce the payment schedule for a larger group.

The right-side of the table documents the result for $N = 5$. It shows that the ratio of average consumption in date 2 to that in date 1 increases the whenever planner must provide incentives. And indeed the necessary increase in this proxy for interest rates, as one moves from weak to strong implementable allocations, is low for $\theta = .5$ and much higher for $\theta = .75$. Provided that correlations in types is a reason for being concerned about runs, the table illustrates a basic result across our experiments: a high level of interest rates is the golden rule for financial stability.

We found a bit surprising that the cost increases with N for these small economies. We have seen, while analyzing Figure 3, that w is much more negative, and shifts cause by population increases are larger when θ is high. This pattern points towards a higher cost and slower convergence under persistence, since patient agents demand more to not run, and becomes relatively more demanding when N increases. On one hand, as with independence, the cost of fulfilling such demands is expected to decrease as population grows and the aggregate state becomes more predictable. On the other hand, while the table shows that the reduction in aggregate uncertainty is sufficient to offset increasing demands under independence, it is not under persistence. But since the uncertainty about the fraction of impatient people in the population should vanish as $N \to \infty$ even when $\theta = .75$ (as N grows the number of Markov switches also grows, generating equal representation of impatient and patient people with increasingly high probability), we expect the cost to reach a maximum at larger N.

5 Insolvency

A fundamental feature of the economies studied until here is that depositor actions are very easy to monitor. We now consider an extension in which the bank is not able to recognize people's identities during the whole first date. We assume that it is possible for some patient people to make two successive accesses to the bank, and that this ability is available with probability $q > 0$. These are the potential 'insiders' that can cause large losses to the financial system in case of a bank run. We also assume that, at the second date, the actions taken at date 1 are matched to the people claiming transfers at the date 2. Hence the identities are recognized and matched to actions at the second date, but this may happen too late to make a difference in case of a run. In summary, the record of actions become updated with a lag, as in Kocherlakota and Wallace (1998). We shall see that the bank has to become more conservative. It has to keep a high level of reserves because some insiders may appear with a positive probability, and in order to keep these insiders away from embezzlement options the bank must increase interest rates with regards to the consumption of this subgroup.

The main purpose of this investigation is to show another side of weak implementation, meaning that stability programs should be taken seriously, using the basic ideas for calculation stability costs outlined above. For simplicity we focus below on some basic properties for this economy under weak implementation and independent shocks. We believe that, at least for small q, the main results about correlation and population increases apply. In what follows, we want to show how imperfect monitoring affects the level of insurance that can be provided, and the corresponding distortions. With this small change in the model, patient people will not be treated equally.[8]. And, we should keep in mind, that if run are possible now (and their existence follows from similar parameter configurations as those seen above) then it is demonstrated that a reasonably small change in the Diamond-Dybvig, following the contributions of Peck and Shell (2003) and EK, can produce cases of insolvency in the sense that a group of people consume zero in equilibrium, thus facing in a very low utility in this scenario.

In the standard model, when actions are fully monitored, this double access assumed here would have no effect on optimal allocations: once date-1 consumption is first transferred to someone, and this person is identified at the second access, the optimum arrangement would trivially give no consumption at the second access. But the situation changes with the imperfect monitoring that we are assuming now. It is true that, under truth-telling, patient people will be

[8] We also find unequal treatment of patient people under weak implementation for the conventional Peck-Shell economy when there is persistence and $p \neq \frac{1}{2}$, for the Green-Lin economy with independence and active constraints, and under strong implementation for Peck-Shell economies with persistence.

lead to consume at date 2, even those with the special abilities. But it is now important to convince the ones with double access to reveal their status freely. Under the current monitoring assumption, the bank is not able to detect this second access and, therefore, must distort allocations if it wants this information revealed.

As mentioned above, we now restrict attention to specifications with independent shocks. The economy can be seen as populated by three types of agents. Type 0 individuals are impatient and have only one access to the bank. A person draws this type with probability $p_0 = p$. Patient agents with single access are designed by type 1. This is drawn with probability $p_1 = (1-p)(1-q)$. Finally, double-access patient people are called type 2. This type occurs with probability $p_2 = (1-q)q$. Since we are interested in computing optimal allocation under truth-telling beliefs, the second access can viewed as just a way for type 2 people to separate themselves from type 1. We shall see that we can leave the issue of double access confined to the introduction of a new truth-telling constraint, which allows us to keep the same structure with N access to the bank as before for much of the accounting that is needed. Accordingly, we set the aggregate state as $\omega \in \{0,1,2\}^N$, which occurs with probability $P(\omega) = p_0^{|\omega|_0} p_1^{|\omega|_1} p_2^{|\omega|_2}$, where $|\omega|_i \equiv \sum_k I_{[\omega_k = i]}$.

A mechanism is (x, y, z), where in addition to impatient date-1 consumption x and date-2 consumption y for type 2, it is introduced date-2 consumption z for type 2. The list (x, y, z) is feasible if

$$\sum_{i=1}^{N} \left(I_{[\omega_i=0]} x_i(\omega^i) + R^{-1} \left[I_{[\omega_i=1]} y_i(\omega) + I_{[\omega_i=2]} z_i(\omega) \right] \right) \leq Y. \tag{9}$$

It is incentive-compatible if

$$E\left[\frac{1}{N} \sum_{i=1}^{N} u\left(y_i\left(\omega_{-i}, 1 \right) \right) \right] \geq E\left[\frac{1}{N} \sum_{i=1}^{N} u\left(x_i\left(\omega_i^-, 0 \right) \right) \right] \tag{10}$$

and

$$E\left[\frac{1}{N} \sum_{i=1}^{N} u\left(z_i\left(\omega_{-i}, 1 \right) \right) \right] \geq E\left[\frac{1}{N} \sum_{i=1}^{N} u\left(x_i\left(\omega_i^-, 0 \right) + x_{i+1}\left(\omega_i^-, 0, 0 \right) \right) \right], \tag{11}$$

where $x_{N+1}(\omega^{N-1}, 0, 0) = 0$. The planner's problem is that of maximizing $U(x, y, z)$, defined as

$$E\left[\frac{1}{N} \sum_{i=1}^{N} \left(I_{[\omega_i=0]} A u\left(x_i\left(\omega_i \right) \right) + I_{[\omega_i=1]} u\left(y_i(\omega) \right) + I_{[\omega_i=0]} u\left(z_i(\omega) \right) \right) \right], \tag{12}$$

subject to (9), (10) and (11).

A computational method similar to that described for environment without insolvency can be designed for this economy. The main difference between the two cases is that the deviation payoff for type-2 individuals are not separable among positions [there are two transfers x_i and x_{i+1} inside the utility function on the right-hand side of (11)]. In this situation the recursive formulation remains valid, but we are not able to get a close solution for the optimal transfers at each position anymore. We propose to guess a solution by approximating first-order conditions and then to iterate on them in order to achieve a numerical convergence to the true solution. The procedure is outlined in the appendix A.

5.1 Numerical findings

We study the existence of pure-strategy run equilibria under the imperfect-monitoring assumption. The following table summarizes the basic results. It presents patient's expected payoff in telling the truth (net of the expected payoff in lying) when he or she believes that all other patient agents (type 1 and type 2) are lying. Parameterization is $N = 5$, $\delta = 2$, $p_0 = 1/2$, $y = 3$, and $R = 1.05$. The table shows the effects of changes in q, the average fraction of insiders among patient individuals, and in A.

A	type	second access probability (q)			
		0.0%	0.5%	1.0%	2.0%
1	1	0.0087	0.0105	0.0123	0.0157
	2	$\cdots$	−0.0299	−0.0296	−0.0290
10	1	−0.0043	−0.0033	−0.0024	−0.0006
	2	$\cdots$	−0.0315	−0.0313	−0.0310

There are eight examples in the table, one for each pair (A, q). Accordingly, first rows (after labels row) refers to a economy in which $A = 1$, and first column (after column *type*) corresponds to iid economy with perfect monitoring (PM) studied in the previous sections. It can be seen that there exists a complete (all patient runs) pure-strategy run only when $A = 10$. Such result is consistent with the PM iid case: there is not insurance enough to sustain a run equilibrium since A, δ, and N are low. This suggests that run equilibria is as easy to find in IM case as in the PM case. However, the fact that run strategy is always much more attractive for type-2 people is evidence that partial runs (in which only type 2 runs) can exist in economies where perfect monitoring would eliminate it. In this sense, a new source of instability emerges. We have seen that insurance level is the essence of run existence in the PM case. Now, a run would exist in economies with not so high insurance level, but with weaker monitoring capacity.

Increasing q from zero to 2% always decreases willingness to run since the relative payoff in telling the truth increases in all rows when changing columns from the left to the right. The reason is that the more probable is type 2, more distortion is necessary to incentive double-access patient agent to tell the truth. Such distortions reduce insurance level and, therefore, the incentive to run.

6 Final remarks

In this paper, we have taken the Peck-Shell (2003) setup as a benchmark for studying the occurrence of runs, and have identified the important role played by the provision of insurance/liquidity on financial stability, with particular emphasis on the size of the population. We have also presented a stronger version of the limit result of Bertolai et al. (2011) in the following sense. We use a generalization of the no-run-constraint concept due to Peck and Shell (2003) and extend the recursive method due to Ennis and Keister (2009) to incorporate the constraint in the case of homothetic preferences. We then show that runs can be eliminated in the standard model at a low cost when the population is small, and at a negligible cost when the population is increased, without appealing to disclosure of information or taking the limit as the population size grows to infinity as in Bertolai et al. (2011).

Having showed that runs are pervasive, and that the most efficient way to avoid them may require a careful computation of the effects of mixed strategies, we also discuss the effects of persistence on the generation of types along the queue, as well as the effects of insolvency. We find that persistence increases the cost of avoiding runs by more than an order of magnitude, and that for small populations the costs are substantially higher in comparison to the independence case. We then show that a small change in the model may produce insolvency. While the level of liquidly insurance falls down as a result of the need to avoid insolvency, bank fragility can still be generated as before. This possibility suggests that the concept of no-run constraints should be taken seriously in a broad sense.

The recursive method has proven very tractable with active constraints. Its construction highlights a simple and enduring message for all experiments: bank stability requires a high interest rate. This golden-rule kind of principle is in contrast with the discussion presented by Wallace (1988) of the main features of the Diamond-Dybvig model. There it is argued that intermediaries would compete to supply contracts with implicit returns resembling those simulated here under weak implementation. The need to assure uniqueness of equilibrium, which was not addressed by Wallace (1988), now tell us that the interest rates supported by this kind of competition might be too low to convince people to focus on the future, a need strongly emphasized by Thornton among other early thinkers.

References

[1] Andolfatto, D., Nosal, E. and Wallace, N. 2007. The role of independency in the Green-Lin Diamond-Dybvig model. *Journal of Economic Theory* 137:709–715.

[2] Bertolai, J., Cavalcanti, R., and Monteiro, P. 2011. A note on convergence of Peck-Shell and Green-Lin mechanisms in the Diamond-Dybvig model. *Working paper.*

[3] Cavalcanti, R. 2010. Inside-money theory after Diamond and Dybvig. *Federal Reserve Bank of Richmond Economic Quarterly* 96(1): 59-82.

[4] Cavalcanti, R. and Monteiro, P. 2011. Enriching information to prevent bank runs. *Working paper.*

[5] Diamond, D. and Dybvig, P. 1983. Bank runs, deposit insurance and liquidity. *Journal of Political Economy* 91:401–19.

[6] Ennis, H. and Keister, T. 2009b. Run equilibria in the Green-Lin Model of financial intermediation. *Journal of Economic Theory* 144:1996–2020.

[7] Green, E. and Lin, P. 2003. Implementing efficient mechanisms in a model of financial intermediation. *Journal of Economic Theory* 109:1–23.

[8] Peck, J. and Shell, K. 2003. Equilibrium bank runs. *Journal of Political Economy* 111:103–123.

[9] Wallace, N. 1988. Another attempt to explain an illiquid banking system: The Diamond-Dybvig model with sequential service taken seriously. *Quarterly Review of the Federal Reserve Bank of Minneapolis* 12(4):3–16.

[10] Wallace, N. 1990. A banking model in which partial suspension is best. *Quarterly Review of the Federal Reserve Bank of Minneapolis* 14(4):11–24.

A The imperfect-monitoring algorithm

As in the perfect monitoring case, we guess the true value for λ, compute optimal solution relative to this guess, and then update multiplier guess in a penalty function fashion. We still calculate relative optimal solution recursively, but now there is an additional procedure in doing so. It is motivated by the non-separability in date 1 consumption in type-2 truth-telling constraint. In what follows, we present how recursive computation of optimal lagrangian value is modified by this non-separability.

First, observe that if we define

$$w_1(x,y) \equiv E\left\{\frac{1}{N}\sum_{i=1}^{N}\Big[u\Big(y_i\left(\omega_{-i},1\right)\Big)-u\Big(x_i\left(\omega_i^-,0\right)\Big)\Big]\right\}$$

and

$$w_2(x,z) \equiv E\left\{\frac{1}{N}\sum_{i=1}^{N}\Big[u\Big(z_i\left(\omega_{-i},2\right)\Big)-u\Big(x_i\left(\omega_i^-,0\right)+x_{i+1}\left(\omega_i^-,0,0\right)\Big)\right\}$$

then the lagrangian in the imperfect monitoring case is

$$\mathcal{L}(x,y,z) = U(x,y,z)+\lambda_1 w_1(x,y)+\lambda_2 w_2(x,z).$$

Proposition 6 *Suppose that we know the set $\theta = \{\theta_1,...,\theta_N\}$, where $\theta_n(\omega^{n-1})$ is the ratio $x_n(\omega^{n-1},0)/x_{n+1}(\omega^{n-1},0,0)$ at the optimal solution relative to the current guess for λ. Then the optimum value for the lagrangian relative to λ is*

$$\mathcal{L}(\lambda) = \frac{1}{N}u(Y)\left(f_1^{\mathbf{0}}(1)\right)^{\delta}$$

where $f_n^{\boldsymbol{j}}(t)$ is defined as

$$\left\{p_0\left[\frac{(\alpha_n^t)^{1-\delta}[a^{\delta}-\lambda_2(p_2/p_0)(1+\theta_n^{-1})^{1-\delta}]+f_{n+1}^{\boldsymbol{j}}(0)}{(\alpha_n^t+f_{n+1}^{\boldsymbol{j}}(0))^{1-\delta}}\right]+\sum_{k=1}^{2}p_k\left\{f_{n+1}^{\boldsymbol{j}_k^+}(k)\right\}^{\delta}\right\}^{\frac{1}{\delta}}$$

in which $\alpha \equiv \left(A-\lambda_1\frac{p_1}{p_0}\right)^{\frac{1}{\delta}}$ and $\alpha_n^t \equiv \left(\alpha^{\delta}-\lambda_2\frac{p_2}{p_0}(1+\theta_n^{-1})^{\delta}-\lambda_2\frac{p_2}{p_0^2}\frac{I_{[t=0]}}{(1+\theta_{n-1})^{\delta}}\right)^{1/\delta}$, and the condition

$$f_{N+1}^{(j_1,j_2)}(\omega) \equiv R^{1/\delta-1}\sum_k j_k(1+\lambda_k)^{1/\delta}$$

Proof. Collecting terms for date 2 consumptions in history ω, and using equal treatment among type 1 and among type 2, we have that planner must choose how to allocate an amount $a(\omega)R$ between patient types in date 2

$$\max\left\{\sum_t (1+\lambda_t)|\omega|_t u\left(\frac{\phi_t a(\omega)R}{|\omega|_t}\right); \phi_1+\phi_2\le 1\right\} =$$
$$u(a(\omega)R)\min\left\{\sum_t (1+\lambda_t)(|\omega|_t)^\delta (\phi_t)^{1-\delta}; \phi_1+\phi_2\le 1\right\}$$

where $|\omega|_t = \sum_i I_{[\omega_i=t]}$. The solution is easily seen to be $\phi_t = \frac{|\omega|_t(1+\lambda_t)^{1/\delta}}{\sum_k |\omega|_k(1+\lambda_k)^{1/\delta}}$, which produces optimal value

$$v_{N+1}(\omega) \equiv u(a(\omega)R)\left\{\sum_k |\omega|_k (1+\lambda_k)^{1/\delta}\right\}^\delta = u(a(\omega))\left\{f_{N+1}^{(|\omega|_1,|\omega|_2)}\right\}^\delta$$

Consider planner choice before the last position and after meeting j_1 patient agents of type 1 and j_2 patient agents of type 2. If $a = a(\omega^{N-2},t)$ denotes the amount in the bank at this moment, then

$$v_N^{\boldsymbol{j}}(\omega^{N-2},t) = \max_{x_N}\left\{p_0\left[\alpha^\delta u(x_N) - \lambda_2\frac{p_2}{p_0}u(x_N+x_{N+1}) + \left\{f_{N+1}^{\boldsymbol{j}}\right\}^\delta u(a-x_N)\right]\right.$$
$$\left. + \sum_k p_k\left[\left\{f_{N+1}^{\boldsymbol{j}_k^+}\right\}^\delta u(a)\right] - I_{[t=0]}\lambda_2\frac{p_2}{p_0}V(x_{N-1}+x_N)\right\}$$

where $\boldsymbol{j} = (j_1,j_2)$, $\boldsymbol{j}_1^+ = \boldsymbol{j}+(1,0)$, and $\boldsymbol{j}_2^+ = \boldsymbol{j}+(0,1)$. Observe that the type of the last agent to access the bank, t, is relevant to determine current consumption. The solution satisfies

$$\left(\frac{a}{x_N}-1\right)^\delta\left[(\alpha_N^1)^\delta - \lambda_2\frac{p_2}{p_0^2}\frac{I_{[t=0]}}{(x_{N-1}/x_N+1)^\delta}\right] = (f_{N+1}^{\boldsymbol{j}})^\delta$$

If we use the ratio θ_{N-1}, the solution is $x_N = \left(1+f_{N+1}^{\boldsymbol{j}}/\alpha_N^t\right)^{-1}a$. Now, plugging these solutions in the objective function we have

$$v_N(\omega^{N-2},t) = u(a)(f_N^{\boldsymbol{j}}(t))^\delta - I_{[t=0]}\lambda_2\frac{p_2}{p_0}u(x_{N-1}+x_N^*)$$

Now, consider planner choice before the last two positions. If we define $v_{N-1}(\omega^{N-3},t)$ as

$$\max_x\left\{p_0\left[\alpha^\delta u(x)\right] + \sum_{k=0}^2 p_k v_N(\omega^{N-2},k) - I_{[t=0]}\lambda_2\frac{p_2}{p_0}u(x_{N-2}+x)\right\}$$
$$= \max_x\left\{p_0\left[\alpha^\delta u(x) + (f_N^{\boldsymbol{j}}(0))^\delta u(a-x) - \lambda_2\frac{p_2}{p_0}u(x+x_N^*)\right] + \right.$$
$$\left. +u(a)\sum_{k=1}^2 p_k\left\{f_N^{\boldsymbol{j}_k^+}(k)\right\}^\delta - I_{[t=0]}\lambda_2\frac{p_2}{p_0}u(x_{N-2}+x)\right\}$$

Necessary first-order condition is

$$\left(\frac{\alpha}{x}\right)^{\delta} - \lambda_2 \frac{p_2}{p_0^2} \frac{I_{[t=0]}}{(x_{N-2}+x)^{\delta}} - \lambda_2 \frac{p_2}{p_0} \frac{1}{(x+x_N^*)^{\delta}} = \frac{(f_N^{\boldsymbol{j}}(0))^{\delta}}{(a-x)^{\delta}}$$

Using the ratios $(\theta_{N-1}, \theta_{N-2})$ and a_{N-1}^t, we can rewrite it as

$$\left(\frac{\alpha_{N-1}}{x}\right)^{\delta} - \lambda_2 \frac{p_2}{p_0} \frac{I_{[t=0]}}{(x_{N-2}+x)^{\delta}} = \frac{(f_N^{\boldsymbol{j}}(0))^{\delta}}{(a-x)^{\delta}}$$

whose solution is $x_{N-1} = (1 + f_N^{\boldsymbol{j}}(0)/\alpha_{N-1}^t)^{-1} s$. Plugging this solution in the objective function, we have

$$v_{N-1}(\omega^{N-3}, t) = u(a) \left\{ f_{N-1}^{\boldsymbol{j}}(t) \right\}^{\delta} - I_{[t=0]} \lambda_2 \frac{p_2}{p_0} u(x_{N-2} + x_{N-1}^*)$$

Similar algebra can be used to get the result

$$v_n(\omega^{n-2}, t) = u(a) \left\{ f_n^{\boldsymbol{j}}(t) \right\}^{\delta} - I_{[t=0]} \lambda_2 \frac{p_2}{p_0} u(x_{n-1} + x_n^*)$$

for all $n < N-1$. The last step is to note that $v_1(\emptyset)$ equals the maximum lagrangian value relative to λ. ■

The previous result shows that if we know θ we are able to compute optimal solution relative to λ by iterating object f_n. However, we generally do not know θ. What we do is to is to guess this set. After computing a solution relative to such guess, we verify if the ratios defined by the current solution and this set agree. If not, we use this new set of ratios as a guess for the next step. This procedure is repeated until convergence. After convergence, we evaluate violation in truth-telling constraints and update multipliers if necessary.

For completeness, we report recursive relation on truth-telling constraints. We have $w_t = \frac{1}{N} u(a) g_1^0(t)$, where $g_1^0(t)$ can be recursively computed using

$$g_n^j(t) = p_0 \left[\frac{g_{n+1}^j(t)(f_{n+1}^{\boldsymbol{j}}(0))^{1-\delta} - \Theta_n(t)(\alpha_n^t)^{1-\delta}}{(\alpha_n^t + f_{n+1}^{\boldsymbol{j}}(0))^{1-\delta}} \right] + \sum_{k=1}^{2} p_k \left[g_{n+1}^{j_k^+}(t) \right]$$

and condition $g_{N+1}^j(t) \equiv j_t \left(\frac{R(1+\lambda_t)^{\frac{1}{\delta}}}{\sum_k j_k (1+\lambda_k)^{\frac{1}{\delta}}} \right)^{1-\delta}$, where $\Theta_n(t) = \frac{p_t}{p_0} \left(1 + \frac{I_{[t=2]}}{\theta_n}\right)^{1-\delta}$. This relation can be obtained by just plugging optimal solution in w_t, from the end of the queue to its beginning.