Discussion Paper Series

RIEB

Kobe University

DP2013-03

Exact Sampling from the Stationary Distribution of Entry-Exit Models

Takashi KAMIHIGASHI

John STACHURSKI

January 23, 2013

Research Institute for Economics and Business Administration

Kobe University

2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Exact Sampling from the Stationary Distribution of Entry-Exit Models*

Takashi Kamihigashi[a] and John Stachurski[b]

[a]*Research Institute for Economics and Business Administration, Kobe University*

[b]*Research School of Economics, Australian National University*

January 23, 2013

Abstract

In equilibrium models of firm dynamics, the stationary equilibrium distribution of firms summarizes the predictions of the model for a given set of primitives. Focusing on Hopenhayn's seminal model of firm dynamics with entry and exit (Econometrica, 60:5, 1992, p. 1127–1150), we provide an algorithm that samples *exactly* from the stationary distribution for any specified exit threshold. The algorithm is able to rapidly generate large numbers of exact and independent draws from this distribution, and can therefore be used to obtain unbiased estimates and confidence intervals for moments and distributions of interest.

Keywords: Simulation, stationary equilibrium, firm dynamics
Journal of Economic Literature Classifications: C61, C63

1 Introduction

The dynamics of entry and exit by firms play an essential role in economic theory as well as in real life. Productive new entrants replace unproductive incumbents, rejuvenating overall economic activity. There is a large and growing literature on this economic mechanism (see, e.g., Hopenhayn and Rogerson, 1993; Cooley and Quadrini, 2001; Melitz, 2003;

*We gratefully acknowledge financial support from ARC Discovery Outstanding Researcher Award DP120100321.

Clementi and Palazzo, 2010), and much of this literature builds upon the model of entry and exit studied by Hopenhayn (1992). The stationary distribution of entry-exit models of the type studied by Hopenhayn represents a cross-sectional distribution of firms that is both consistent with the definition of equilibrium at any point in time and also invariant over time. For typical parameter values the stationary distribution is uniquely defined but has no analytical solution. To compare the model with data, it is therefore necessary to either numerically compute the distribution or simulate from it.

Direct computation of the distribution involves discretization of some form. The size of the resulting errors are typically unknown, reducing confidence in quantitative results. Regarding simulation, for a given exit policy it is straightforward to set up an ergodic Markov process such that its stationary distribution coincides with the cross-sectional stationary distribution, and this permits approximate sampling from the stationary distribution simply by running the process from an arbitrary initial value until it is judged to have "nearly" converged. Simulating until the distribution of the state is approximately stationary is referred to as "burn-in." Unfortunately, the length of burn-in required is the subject of guesswork and heuristics, which is clearly a poor foundation for analysis. Moreover, regardless of how much burn-in is performed, the resulting sample is never exactly stationary, and the size of the error is once again unknown. Finally, for a given method, the size of the error is likely to vary with the parameters, since the parameters change the structure of the problem. If the burn-in is not varied accordingly, this is likely to cause bias.

In this paper we show that these problems can in fact be overcome. By using a variation of Propp and Wilson's (1996) coupling from the past (CFTP) technique, we show that it is possible to sample exactly from the stationary distribution of this class of models, for any specified exit threshold. In particular, we develop an efficient algorithm that generates exact, IID draws from the stationary distribution. For each random seed, the algorithm terminates as soon as an exact draw has been generated, and it is guaranteed to terminate in finite time with probability one. Hence there is no need for the heuristics used to judge burn-in time. Moreover, by repeating the algorithm with independent seeds it becomes possible to generate multiple independent draws from the stationary distribution. These processes can be run sequentially or in parallel. Even with sequential generation, the algorithm permits rapid generation of exact IID draws from the stationary distribution. Because the resulting sample is IID and exact, our algorithm can be used to obtain *unbiased* estimates and confidence intervals for moments and distributions of interest.

The CFTP technique has been used successfully for models with large but discrete state spaces, where standard methods for computing stationary distributions are infeasible. Applications range from statistical mechanics to page ranking and the design of peer-to-peer file sharing systems (see chapter 2 of Levin *et al.*, 2008 for recent survey). Unfortunately, in the case of *continuous* state spaces, applicability of CFTP is not automatic. It depends on the particular structure of the model in question, and the ingenuity required to exploit it. Murdoch and Green (1998) showed that CFTP can in principle be used in continuous state settings when the underlying Markov process satisfies Doeblin's condition. Doeblin's condition is a technical condition requiring the existence of a nonnegative function that (a) integrates to a positive value, (b) depends only on the next state, and (c) is pointwise dominated by the transition density function (which depends on both the current state and the next). However, Murdoch and Green (1998, p. 48) admit that their basic method, which is in principle applicable to our model satisfying Doeblin's condition, may have "a limited range of application for two reasons." The first reasons is the function associated with Doeblin's condition "may be too small for practical use" to generate exact draws in a reasonable length of time. Second, their method "requires the user to know" the transition density of the underlying Markov process even though it is can be nontrivial or even impossible to explictly calculate this density.[1]

Another setting where CFTP methods have met with success in continuous state applications is when the underlying Markov process is monotone. For such processes, the CFTP method can be highly efficient. Corcoran and Tweedie (2001) developed general results on CFTP particularly suitable for monotone Markov processes, and in the economic literature, Nishimura and Stachurski (2010) applied their ideas to an incomplete markets model. Unfortuntately, the CFTP methods for monotone processes are not applicable to models of firm dynamics, which are typically not monotone. In fact, possile non-monotonicity is an important feature of the dynamics of entry and exit: the least productive firm in the market today is likely to be less productive than the average new entrant tomorrow.

Unlike the CFTP methods used by Nishimura and Stachurski (2010), the algorithm we develop in this paper is entirely new, and is easily implementable even though our model is not monotone. The state space is continuous, and, while the Markov process for productivity of incumbents is monotone increasing, the overall process that combines dynamics of incumbents and new entrants is not. We show that this algorithm successfully terminates with probability one by exploiting both the monotonicity of productivity for incumbents and the renewal property introduced by new entrants.

Our algorithm is clearly distinct from Murdoch and Green's (1998) method discussed above, in that it does not use Doeblin's condition, and does not require the explicit knowledge of the transition density.[2] It is therefore unnecessary to find a function associated with Doeblin's condition, or to explicitly calculate the transition density. As long as one can simulate the overall Markov process, one can sample exactly from the stationary dis-

[1] See Foss and Tweedie (1998) and Athreya and Stenflo (2003) for theoretical extensions of Murdoch and Green's (1998) work.

[2] The assumptions used to show the probability one termination of the algorithm in fact imply Doeblin's condition for some n-step transition, but our proof of this property does not use the latter.

tribution using our algorithm.[3]

The rest of the paper is organized as follows. In section 2 we offer an intuitive explanation of the CFTP technique, briefly describe Hopenhayn's (1992) entry-exit model, and introduce a Markov process charactering firm dynamics in preparation for exact sampling. In section 3 we develop our algorithm, show that it successfully terminates with probability one, and apply it to two numerical specifications of the model.

2 Preliminaries

In this section we briefly review CFTP, and recall the major features of Hopenhayn's (1992) model.

2.1 Coupling From the Past

There are many versions of the CFTP algorithm. We now review a version for finite state Markov chains that provides a useful stepping stone to understanding the algorithm developed below. To begin, consider an irreducible aperiodic Markov chain $\{X_t\}$ on finite state space S with unique stationary distribution μ. Suppose that the dynamics of the model are given by $X_{t+1} = h(X_t, W_{t+1})$ where $\{W_t\}$ is IID. If we start at any X_0 and simulate $\{X_t\}$ by drawing $W_1, W_2, \ldots$ and successively applying the maps $h(\cdot, W_1), h(\cdot, W_2), \ldots$, then the distribution of X_t will converge to (but in general never reach) the stationary distribution μ. A second immediate observation is that if $\{X_t\}$ and $\{X'_t\}$ are two chains simulated using the *same sequence* of shocks $\{W_t\}$, and if $X_\tau = X'_\tau$ for some τ, then $X_t = X'_t$ for all $t \geq \tau$. We say that $\{X_t\}$ and $\{X'_t\}$ *couple* at time τ.

Now consider an IID sequence of shocks $\{W_t\}_{t=-\infty}^{0}$ indexed on $\{\ldots, -2, -1, 0\}$, and let $\{X_t\}_{t=-\infty}^{0}$ be the process generated from these shocks, starting in the infinite past and terminating at $t = 0$. Intuitively, since the burn-in is infinite, the distribution of X_0 will be exactly μ. On the other hand, because the chain $\{X_t\}_{t=-\infty}^{0}$ is infinite and hence cannot be simulated, it appears that X_0 is not observable. In fact this is not the case. To understand how X_0 can be observed, suppose that we fix $T \in \mathbb{N}$, take the finite subset of shocks $W_{-T+1}, W_{-T+2}, \ldots, W_0$ and then, for each point x in the state space S, construct one version of the process $\{X_t(x)\}_{t=-T}^{0}$ that starts at this point x at time $-T$ and runs up until time zero. All of these processes $\{X_t(x)\}_{t=-T}^{0}$ are updated using the same shocks $W_{-T+1}, W_{T+2}, \ldots, W_0$. We will call them *tracking processes*. Because there are only finitely many tracking processes and each is of finite length, it is possible to simulate them. Two different visualizations of the tracking processes are shown in the top panel of figures 1

[3] Exactness is modulo the errors associated with floating point arithmetic, which cannot be avoided.

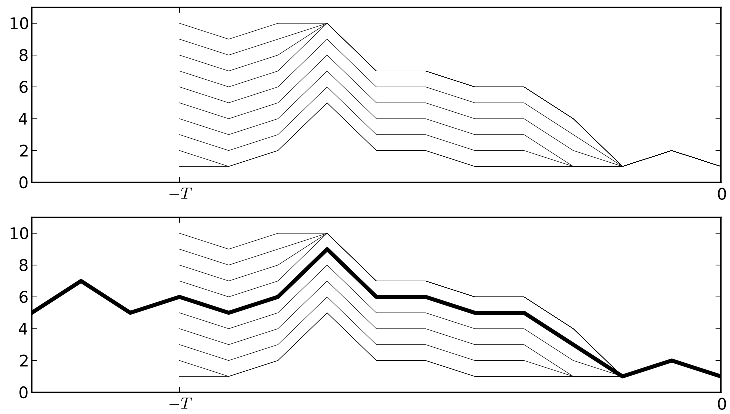

Figure 1: Coalescence of the tracking processes

and 2 respectively. Here the state space is $\{1, \ldots, 10\}$, and the two figures show different realizations from the same model. Notice that in the simulations, some of the processes couple and then run together up until time zero. In figure 1, all of the tracking processes have coupled by time zero, and only one final value is observed. If this is the case we will say that the tracking processes *coalesce*. In figure 2 multiple paths are still present at time zero, and the processes fail to coalesce.

Now consider again the process $\{X_t\}_{t=-\infty}^{0}$ generated from the entire sequence of shocks $\{W_t\}_{t=-\infty}^{0}$. Since this process must pass through one point x in the state space at $-T$, and since it receives the same shocks $W_{-T+1}, W_{-T+2}, \ldots, W_0$ as the tracking processes from that time forwards, over the period $-T, -T+1, \ldots, 0$ it must follow the same path as the tracking process $\{X_t(x)\}_{t=-T}^{0}$ that started at x. Of course we do not know which of the tracking processes it is following. However, *if* the tracking processes coalesce, then there is only *one* final observation. This observation must be a realization of the time zero value X_0 of the process $\{X_t\}_{t=-\infty}^{0}$ that started in the infinite past, and is therefore a draw from the stationary distribution μ. Such an outcome is illustrated in the bottom panel of figure 1. For comparison, an unsuccessful outcome is illustrated in the bottom panel of figure 2. Here there are three final values, and we do not know which is the time zero value of the process $\{X_t\}_{t=-\infty}^{0}$. In this case we can take one step further back in time, drawing the additional shock W_{-T} while preserving the existing shocks $W_{-T+1}, W_{-T+2}, \ldots, W_0$, recalculate the tracking processes, test for coalescence, and so on. This procedure will eventually terminate with an exact draw from μ.[4]

While this technique works when there are a finite number of states, it is clearly not pos-

[4]Using finiteness of the state space, aperiodicity and irreducibility, it can be shown that termination occurs in finite time with probability one.

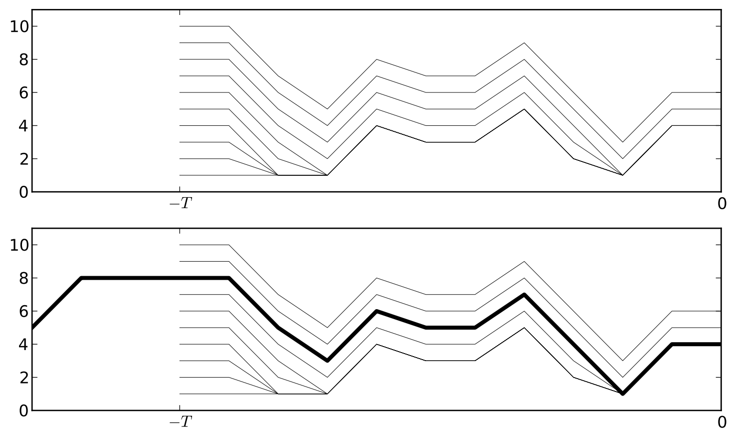

Figure 2: The tracking processes failing to coalesce

sible to directly implement the same idea when the state space is infinite. However, variations on the idea can potentially be found by exploiting the structure of a given model. In the case of the entry-exit model studied below, we show how the monotonicity of productivity for incumbent firms and the renewal caused by new entrants can be combined to pin down the final value X_0.

2.2 The Model

Hopenhayn (1992) developed what is now a benchmark model for studying the evolution of a competitive industry where entry and exit is endogenously determined. In this section we briefly outline the model and review standard (non-exact) simulation. Most of our notation follows Hopenhayn (1992). If X is a random variable, then $\mathcal{D}(X)$ denotes the distribution of X. The symbol $\mathscr{B}$ represents the Borel sets on $[0, 1]$.

In the model there is a continuum of firms that produce a homogeneous good. The firms face idiosyncratic productivity shocks that follow a Markov process on $S := [0, 1]$. The conditional cumulative distribution function for the shock process is denoted by $F(\phi' \mid \phi)$. Following Hopenhayn (1992, assumptions A.3 and A.4), we impose the following restrictions:

Assumption 2.1. F is decreasing in its second argument and, for any $\epsilon > 0$ and any $\phi \in S$, there exists an integer n such that $F^n(\epsilon \mid \phi) > 0$.[5]

[5] $F^n(\cdot \mid \phi)$ is the conditional distribution for productivity after n periods, given current productivity ϕ.

We let P denote the corresponding stochastic kernel. That is, $P(\phi, A) := \int_A F(d\phi' \,|\, \phi)$ for $\phi \in S$ and $A \in \mathscr{B}$. Incumbent firms exit the industry whenever their current productivity falls below a reservation value x_t. Letting M_t be the mass of entrants at time t and ν be the Borel probability measure from which the productivity of entrants is drawn, the sequence of firm distributions $\{\mu_t\}$ on S satisfies $\mu_{t+1}(A) = \int P(\phi, A)\mathbb{1}\{\phi \geq x_t\}\mu_t(d\phi) + M_{t+1}\nu(A)$ for all $A \in \mathscr{B}$, where $\mathbb{1}\{\cdot\}$ is an indicator function. At the stationary equilibrium, both x and M are constant, and a stationary distribution μ is a Borel probability[6] measure μ satisfying

$$\mu(A) = \int P(\phi, A)\mathbb{1}\{\phi \geq x\}\mu(d\phi) + M\nu(A) \qquad (A \in \mathscr{B}). \tag{1}$$

This is eq. (12) in Hopenhayn (1992). It follows from (1) and $\mu(S) = P(\phi, S) = \nu(S) = 1$ that $M = M(x, \mu) := \mu\{\phi \in S : \phi < x\}$. As a result, we can also write (1) as

$$\mu(A) = \int Q(\phi, A)\mu(d\phi) \tag{2}$$

where

$$Q(\phi, A) := P(\phi, A)\mathbb{1}\{\phi \geq x\} + \nu(A)\mathbb{1}\{\phi < x\}. \tag{3}$$

Equation (2) states that μ is a stationary distribution for the stochastic kernel Q in the usual sense of time invariance. As shown by Hopenhayn (1992, p. 13), the kernel Q has only one stationary distribution. For the purposes of this paper we will treat x as given.

2.3 Simulation

Our initial task is to produce an ergodic Markov process the stationary distribution of which coincides with the distribution μ in (2), taking x as a given constant in $(0, 1)$. To do this, we need a method for sampling from the stochastic kernel Q. The first step is to simulate from the conditional distribution $P(\phi, \cdot) = F(\cdot \,|\, \phi)$. In particular, we seek a random variable U and a function g such that $\mathcal{D}(g(\phi, U)) = F(\cdot \,|\, \phi)$ for all $\phi \in S$. This can be achieved via the inverse transform method, where U is uniform on $[0, 1]$ and $g(\phi, u) = F^{-1}(u \,|\, \phi)$.[7] Now consider the process $\{\Phi_t\}$ defined by

$$\Phi_{t+1} = g(\Phi_t, U_{t+1})\mathbb{1}\{\Phi_t \geq x\} + Z_{t+1}\mathbb{1}\{\Phi_t < x\} \tag{4}$$

where $\{(U_t, Z_t)\}$ is IID with $\mathcal{D}(Z_t) = \nu$ and $\mathcal{D}(U_t) = \text{Uniform}[0, 1]$. Comparing (3) and (4), it can be seen that $\{\Phi_t\}$ is a Markov process with stochastic kernel Q.

[6] We focus only on normalized measures, since other cases are just scalar multiples.

[7] Here $F^{-1}(\cdot \,|\, \phi)$ is the generalized inverse of $F(\cdot|\phi)$. That is, $F^{-1}(u \,|\, \phi) := \inf\{z : F(z \,|\, \phi) \geq u\}$.

3 Exact Sampling

Let $\{(U_t, Z_t)\}_{t\in\mathbb{Z}}$ be an infinite sequence of IID shocks indexed on $\mathbb{Z}$ and with each pair (U_t, Z_t) having the product distribution Uniform$[0,1] \times \nu$. To simplify notation we will let $g_t := g(\cdot, U_t)$, so that, for example, $g_t \cdots g_1 \phi := g_t \circ g_{t-1} \circ \cdots \circ g_1(\phi)$ is exogenous productivity at t, given time zero productivity $\phi \in S$. To further simplify notation, let

$$h_t(\phi) := g(\phi, U_t)\mathbb{1}\{\phi \geq x\} + Z_t\mathbb{1}\{\phi < x\},$$

so that (4) becomes $\Phi_{t+1} = h_{t+1}\,\Phi_t := h_{t+1}(\Phi_t)$.

In section 2.1 we discussed CFTP for processes with finite state spaces. Since the number of tracking processes is equal to the number of points in the state space, models with finite state space have only finitely many tracking processes. In the current setting, however, the state space S is a continuum, and hence the number of tracking processes is likewise a continuum. As a result, both occurrence *and* detection of coalescence becomes a considerably more difficult problem.

To describe the tracking processes, fix $T \geq 1$. For each $\phi \in S$, there is a corresponding tracking process that starts at time $-T$ with value ϕ, and then updates with maps $h_{-T+1}, h_{-T+2}, \ldots, h_0$, obtaining the value $h_0 \cdots h_{-T+1}\,\phi$ at time zero. We say that the tracking processes *coalesce* if, for some $T \in \mathbb{N}$, the set of final values

$$h_0 \cdots h_{-T+1}(S) := \{h_0 \cdots h_{-T+1}\,\phi \;:\; \phi \in S\} \tag{5}$$

is a singleton. What we will now show is that under mild conditions coalescence occurs with probability one, and, moreover, that it is not necessary to keep track of the full continuum of tracking processes in order to find the value of the singleton. In particular, we show that, conditional on a certain event described below, the set of final values $h_0 \cdots h_{-T+1}(S)$ has only *finitely* many possibilities. Hence coalescence occurs whenever these finite possibilities take the same value. All of these finite possibilities are computable. To begin describing them, let $T > 1$ be given, let

$$\Sigma_T := \{k \in \mathbb{N} : 1 \leq k < T \text{ and } g_{-T+k} \cdots g_{-T+2} \cdot g_{-T+1}\, 1 < x\},$$

and let $\sigma_T := \min \Sigma_T$. Intuitively, σ_T is the number of periods that an incumbent firm survives, given that it starts at time $-T$ with maximal productivity 1 and faces the shock sequence $\{U_t\}_{t\in\mathbb{Z}}$. Clearly σ_T is only defined when Σ_T is nonempty. However, the probability that Σ_T is nonempty converges to one as $T \to \infty$ by assumption 2.1. Moreover, it is remarkable that if Σ_T is nonempty, then the set $h_0 \cdots h_{-T+1}(S)$, which contains the final values of the tracking processes started at $-T$, can have only finitely many values:

Lemma 3.1. *If Σ_T is nonempty, then $h_0 \cdots h_{-T+1}(S) \subset \Lambda_T$, where*

$$\Lambda_T := \{h_0 \cdots h_{-T+k+1}\, Z_{-T+k} \;:\; k = 1, \ldots, \sigma_T + 1\}. \tag{6}$$

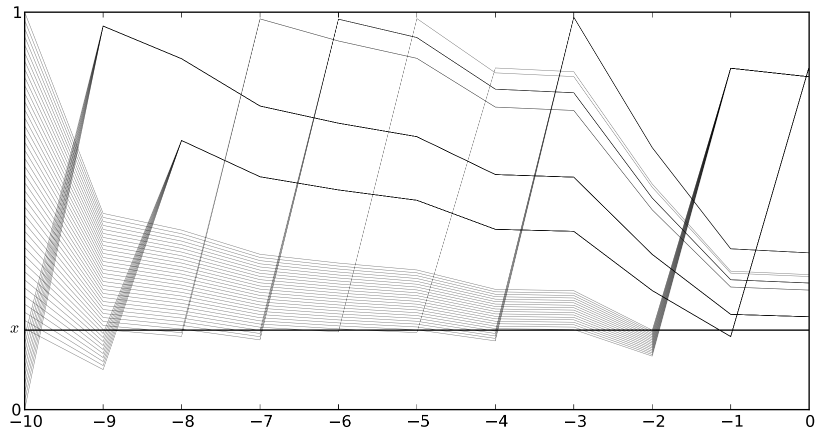

Figure 3: Tracking processes with $T = 10$ and $\sigma_T = 8$

The proof of lemma 3.1 is given in section 5. The intuition for the result can be obtained by considering figure 3. In the figure, $T = 10$. Tracking processes are plotted for 50 different initial values of $\phi \in S$. (Ideally, tracking processes would be plotted from every $\phi \in S$, but this is clearly impossible.) For this particular realization of shocks, the set Σ_T is nonempty because the process starting from 1 at time -10 falls below x at $t = -2$ (and hence $\sigma_T = 10 - 2 = 8$). As is clear from the figure, the fact that the process starting from 1 at time -10 falls below x at $t = -2$ implies that *all* tracking processes fall below x at least once between -10 and -2 (recall that the productivity of incumbents is monotone). Moreover, if any collection of tracking processes fall below x at some point in time t, they subsequently couple, taking the common value Z_{t+1} at $t + 1$ and being equal from then on. As a result, by $t = -1$, there are at most $\sigma_T + 1 = 9$ distinct tracking processes. Their time zero values are included in the set Λ_T defined in lemma 3.1. In particular, Λ_T is the time zero values of the processes that start below x at dates $-10, -9, \ldots, -2$.

To see the importance of lemma 3.1, let $\{\Phi_t\}_{t \in \mathbb{Z}}$ be a stationary, doubly-indexed process on the same probability space as $\{(U_t, Z_t)\}_{t \in \mathbb{Z}}$ that obeys $\Phi_{t+1} = h_{t+1} \Phi_t$ for all $t \in \mathbb{Z}$. The common marginal distribution of Φ_t is μ. Since Φ_{-T} lies somewhere in S, we know that $\Phi_0 = h_0 \cdots h_{-T+1} \Phi_{-T} \in h_0 \cdots h_{-T+1} (S)$. Moreover, if the set Σ_T is nonempty, then lemma 3.1 yields the inclusion $h_0 \cdots h_{-T+1} (S) \subset \Lambda_T$, and Φ_0 lies in the finite observable set Λ_T. In particular, if Λ_T is a singleton, then the value of Φ_0 is revealed as the value of that singleton.

Figures 4 and 5 show simulations with successful and unsuccessful coalescence respectively. In each figure, the top panel shows only the tracking processes. (As with figure 3, the full continuum of tracking processes cannot be plotted, so we show only 50.) The bottom panel shows the tracking processes and the path of $\{\Phi_t\}$. In reality, the path of

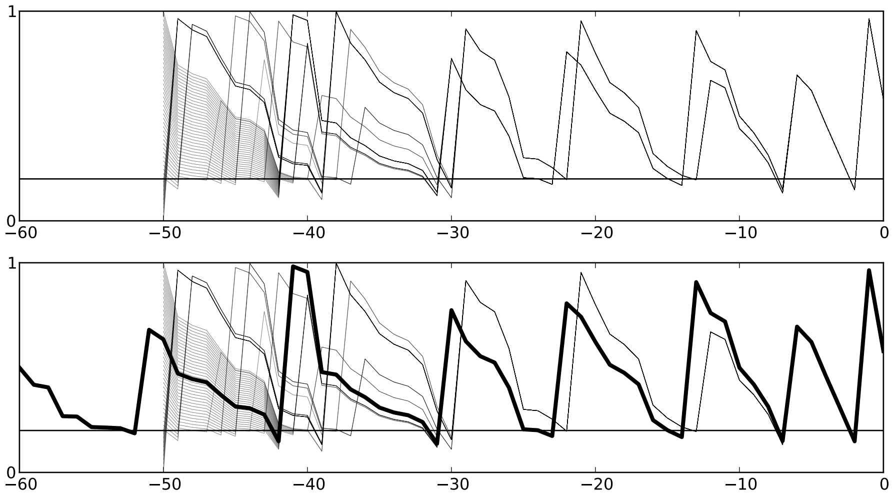

Figure 4: Successful coalescence from $T = 50$

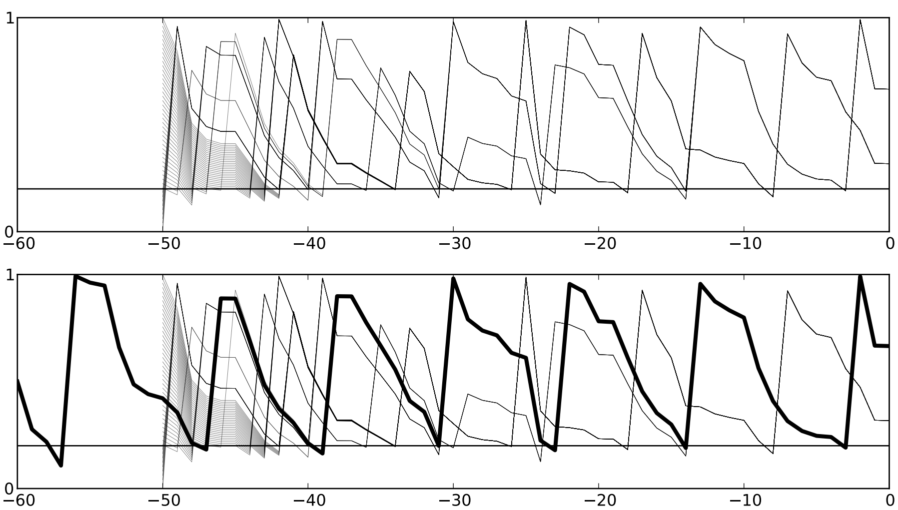

Figure 5: Failure of coalescence

$\{\Phi_t\}$ is not observed. However, in figure 4, there is only one final, coalesced value, and Φ_0 must take this value. Hence Φ_0 is observed. On the other hand, in figure 5, Φ_0 is equal to one of two final values, and we have no way of identifying which one it is.

Now let us consider how to use our results to sample from μ by generating observations of Φ_0. In order to avoid conditioning on coalescence by a certain point in time, we wish to reveal the value of Φ_0 for every random seed. This can be done by fixing the seed, which determines the values of the shock processes, and then taking T larger and larger until coalescence occurs. Algorithm 1 gives details. The algorithm terminates with an exact draw from μ. Replication with independent shocks will generate independent draws.

Algorithm 1: Generates an exact draw from μ

fix T to be an integer greater than 1;
draw $(U_0, Z_0), \dots, (U_{-T+1}, Z_{-T+1})$ independently from their distributions;
repeat
    compute the set Σ_T ;
    if Σ_T *is nonempty* **then**
        compute the set Λ_T ;
        if Λ_T *is a singleton* **then**
            set Φ_0 to be the value of that singleton ;
            break ;
        end
    end
    draw (U_{-T}, Z_{-T}) and append to list $(U_0, Z_0), \dots, (U_{-T+1}, Z_{-T+1})$;
    set $T = T + 1$;
end
return Φ_0 ;

At this stage we do not know that the algorithm will terminate with probability one. This issue is central to the correctness of the algorithm because, as discussed above, the way we avoid conditioning is by revealing the value of Φ_0 for every random seed. We now show that probability one termination in finite time holds under the following condition, which is satisfied by many standard distributions.

Assumption 3.1. If $G \subset S$ is a nonempty open set, then $\nu(G) > 0$.

Proposition 3.1. *Let $T^* := \min\{T \in \mathbb{N} : \Sigma_T$ is nonempty and Λ_T is a singleton $\}$. If assumption 3.1 holds, then there exists a $\gamma \in (0, 1)$ such that $\mathbb{P}\{T^* > t\} \leq \gamma^t$. In particular, $\mathbb{P}\{T^* < \infty\} = 1$.*

Note that proposition 3.1 not only gives probability one occurrence, but also provides the geometric rate $\mathbb{P}\{T^* > t\} = O(\gamma^t)$. The proof of proposition 3.1 is given in section 5.

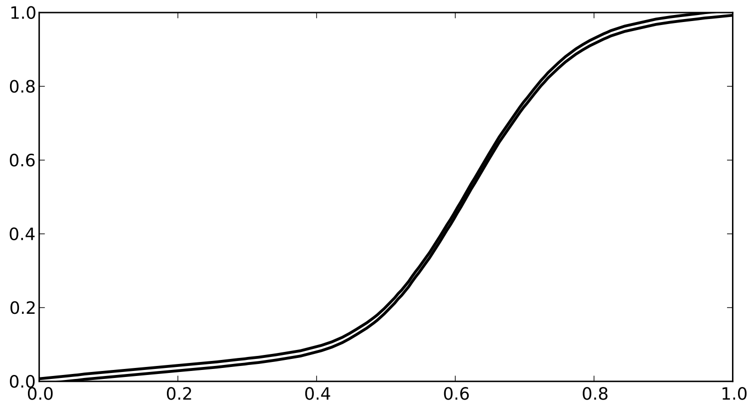

Figure 6: 95% confidence set for the stationary distribution

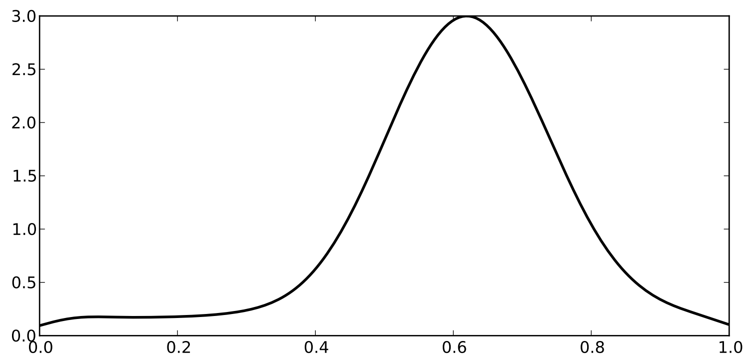

Figure 7: Stationary density, first parameterization

The web page `http://johnstachurski.net/papers/hh_sampling.html` contains a simple implementation of algorithm 1. We tested the code with two different parameterizations. First, we followed Hopenhayn and Rogerson (1993) in taking the distribution ν for new entrants to be uniform, and the process for incumbents to be $\Phi_{t+1} = g_{t+1}\Phi_t = a + \rho\Phi_t + \epsilon_{t+1}$ where $\{\epsilon_t\}$ is IID with distribution $N(0, \sigma^2)$.[8] The parameters were set to $a = 0.36$, $\rho = 0.4$ and $\sigma = 0.1$, while x was set to 0.49, so that approximately 40% of incumbents exit within 5 years (see Hopenhayn, 1992, p. 1127). For these parameters, running the program on our computer produces about 36,000 independent draws from μ per second.[9]

Figure 6 shows a 95% confidence set for the cumulative distribution function corresponding to μ, based on 36,000 observations and calculated using the Kolmogorov distribution of the sup norm deviation between true and empirical cdfs. The Kolmogorov result is

[8] To bound the process we added reflecting barriers at 0 and 1.

[9] The computer is a standard workstation with a 2.67GHz Intel CPU and 4 gigabytes of RAM.

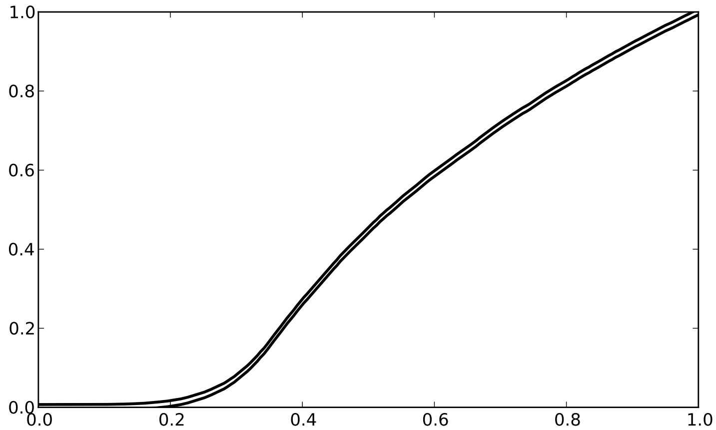

Figure 8: 95% confidence set, second parameterization

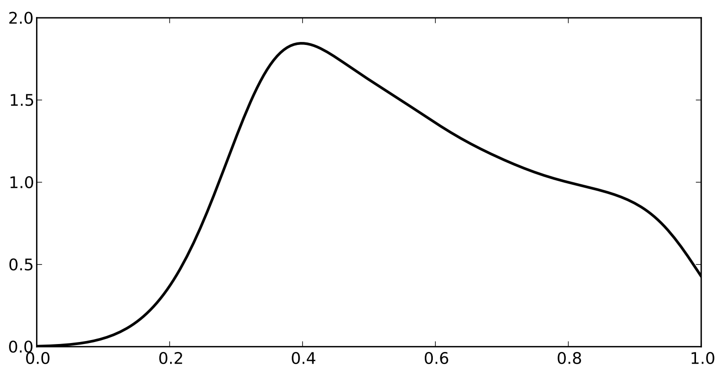

Figure 9: Stationary density, second parameterization

applicable here because the draws are exact and IID. The true distribution function lies entirely between the two bands with 95% probability. Figure 7 shows the corresponding density, computed from the same observations combined with a standard nonparametric kernel density estimator (using a Gaussian kernel).

For the second parameterization we took the process for incumbents to be $\Phi_{t+1} = \Phi_t \epsilon_{t+1}$, where $\epsilon_t \sim \text{Uniform}(\alpha, 1)$, and $\nu = \text{Beta}(5, 1)$. We set α to 0.65 and x to 0.35 to match the same 40% exit rate over 5 years. For these parameters, the program produces about 32,500 independent draws from μ per second. Figure 8 shows a 95% confidence set computed from 32,500 observations for the cumulative distribution function, and figure 9 shows the stationary density calculated with a standard nonparametric kernel density estimator.

As a final illustration of the algorithm, suppose now that we wish to calculate aggregate output $\int f(\phi, \ell)\mu(d\phi)$, where f is a production function and ℓ is given. Letting $f(\phi, \ell) = \phi\, \ell^\theta$ with $\ell = 0.5$ and $\theta = 0.64$, and adopting the first set of parameters, we computed the empirical estimate $Q_n := \frac{1}{n}\sum_{i=1}^n f(\Phi_i, \ell)$ using 36,000 observations (generated in one second). The mean Q_n was 0.3848, while the standard error computed from the CLT was 0.00085, giving a 95% confidence interval for $\int f(\phi, \ell)\mu(d\phi)$ of $[0.3840, 0.3857]$. With more draws the standard errors fall at rate $O(n^{-1/2})$.

4 Conclusion

In this paper we developed an algorithm that samples exactly from the stationary distribution of Hopenhayn's (1992) entry-exit model for a given exit policy. In the applications we considered, the algorithm produced in the order of 30,000 to 36,000 independent exact draws from the stationary distribution per second. The basic idea should extend to many economic applications involving entry and exit.

5 Proofs

In the following proofs we write σ for σ_T to simplify notation.

Proof of lemma 3.1. Let Σ_T be nonempty. As a first step, we show that if $\phi \geq x$, then there exists a $j \in \{1, \ldots, \sigma\}$ such that $h_{-T+j} \cdots h_{-T+1}\, \phi < x$. To see that this is so, fix $\phi \geq x$ and suppose that the statement fails. In other words, $h_{-T+j} \cdots h_{-T+1}\, \phi \geq x$ for $j \in \{1, \ldots, \sigma\}$. We know that if $y \geq x$, then $h_i\, y = g_i\, y$. It follows that $h_{-T+\sigma} \cdots h_{-T+1}\, \phi = g_{-T+\sigma} \cdots g_{-T+1}\, \phi$. But then

$$x \leq h_{-T+\sigma} \cdots h_{-T+1}\, \phi = g_{-T+\sigma} \cdots g_{-T+1}\, \phi \leq g_{-T+\sigma} \cdots g_{-T+1}\, 1 < x,$$

where the second inequality is due to monotonicity of g_i, and then third is by the definition of σ. Contradiction.

To complete the proof, pick any $\phi \in S$. Our claim is that $h_0 \cdots h_{-T+1}\, \phi \in \Lambda_T$. Suppose first that $\phi < x$. In this case we have $h_0 \cdots h_{-T+1}\, \phi = h_0 \cdots h_{-T+2}\, Z_{-T+1}$, which is an element of Λ_T. Next, suppose that $\phi \geq x$. In light of the preceding argument, there exists a $j \in \{0, \ldots, \sigma\}$ with $h_{-T+j} \cdots h_{-T+1}\, \phi < x$, and hence

$$h_{-T+j+1} \cdot h_{-T+j} \cdots h_{-T+1}\, \phi = Z_{-T+j+1},$$

from which we obtain

$$\begin{aligned} h_0 \cdots h_{-T+1}\, \phi &= h_0 \cdots h_{-T+j+2} \cdot h_{-T+j+1} \cdot h_{-T+j} \cdots h_{-T+1}\, \phi \\ &= h_0 \cdots h_{-T+j+2}\, Z_{-T+j+1}. \end{aligned}$$

Since $j \in \{0, \dots, \sigma\}$, the right-hand side is an element of Λ_T. This completes the proof. $\square$

Proof of proposition 3.1. Let n be an integer such that $F^n(x \mid 1) > 0$, existence of which is due to assumption 2.1. Fixing $j \in \mathbb{N}$, let

$$E_j := \{g_{-(j-1)n-1} \cdots g_{-jn} \, 1 < x\} \cap \{Z_{-(j-1)n-1} < x, \dots, Z_{-jn} < x\}.$$

The events $\{g_{-(j-1)n-1} \cdots g_{-jn} \, 1 < x\}$ and $\{Z_{-(j-1)n-1} < x, \dots, Z_{-jn} < x\}$ are independent because the first event depends only on $U_{-(j-1)n-1}, \dots, U_{-jn}$ and the second depends only on $Z_{-(j-1)n-1}, \dots, Z_{-jn}$. As a result,

$$\delta := \mathbb{P}(E_j) = F^n(x \mid 1) \nu([0, x))^n.$$

The constant δ is strictly positive as a result of assumption 3.1. We claim that if the event E_j occurs, then Σ_{jn+1} is nonempty and Λ_{jn+1} is a singleton. To simplify notation, we treat only the case of $j = 1$.

So suppose that E_1 occurs. Clearly Σ_{n+1} contains n, and hence is nonempty. To see that Λ_{n+1} is a singleton, observe that since $\sigma = \sigma_{n+1}$ is the smallest element of Σ_{n+1}, we must have $\sigma \leq n$. As a consequence,

$$\begin{aligned} \Lambda_{n+1} &= \{h_0 \cdots h_{-n+k} \, Z_{-n-1+k} \; : \; k = 1, \dots, \sigma + 1\} \\ &\subset \{h_0 \cdots h_{-n+k} \, Z_{-n-1+k} \; : \; k = 1, \dots, n + 1\}. \end{aligned}$$

We claim that on the set E_1 we have

$$h_0 \cdots h_{-n+k} \, Z_{-n+k-1} = Z_0 \quad \text{for any} \quad k \in \{1, \dots, n+1\}. \tag{7}$$

To prove that (7) holds, observe that on E_1 the values $Z_{-1}, \dots, Z_{-n}$ are all less than x. As a result, we have

$$\begin{aligned} h_0 \, Z_{-1} &= Z_0 \\ h_0 \, h_{-1} \, Z_{-2} &= h_0 \, Z_{-1} = Z_0 \\ h_0 \, h_{-1} \, h_{-2} \, Z_{-3} &= h_0 \, h_{-1} \, Z_{-2} = h_0 \, Z_{-1} = Z_0 \end{aligned}$$

and so on. Together, these equalities give (7). As a consequence, we conclude that Σ_{n+1} is nonempty and Λ_{n+1} is a singleton whenever E_1 occurs, and, more generally, Σ_{nj+1} is nonempty and Λ_{nj+1} is a singleton whenever E_j occurs. The events $E_1, E_2, \dots$ are independent and have positive probability δ. Using the definition of T^*, we then have

$$\mathbb{P}\{T^* > nj\} = \mathbb{P}\{T^* \geq nj + 1\} \leq \mathbb{P} \cup_{i=1}^{j} E_i^c = (1 - \delta)^j$$

for all $k \in \mathbb{N}$. Setting $\gamma := (1 - \delta)^{1/n}$ gives the result stated in the proposition. $\square$

References

[1] Athreya, K. B., Stenflo, O., 2003. Perfect sampling for Doeblin chains. Sankhya 65, 763–777.

[2] Clementi, G. L. and B. Palazzo, 2010. Entry, exit, firm dynamics, and aggregate fluctuations. Mimeo.

[3] Cooley, T.F. and Quadrini, V., 2001. Financial markets and firm dynamics. American Economic Review, 91 (5), 1286–1310.

[4] Corcoran, J. N., Tweedie, R. L., 2001. Perfect sampling of ergodic Harris chains. Annals of Applied Probability 11, 438–451.

[5] Foss, S., Tweedie, R. L., 1998. Perfect simulation and backward coupling. Stochastic Models 14, 187–204.

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

[7] Hopenhayn, H. and Rogerson, R., 1993. Job turnover and policy evaluation: A general equilibrium analysis. Journal of Political Economy, 101 (5), 915–938.

[8] Levin, D. A., Y. Peres, E. L. Wilmer, 2008. Markov Chains and Mixing Times. American Mathematical Society.

[9] Melitz, M.J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica, 71 (6), 1695–1725.

[10] Meyn, S., Tweedie, R. L., 1993. Markov Chains and Stochastic Stability. Springer-Verlag, New York.

[11] Murdoch, D. J., Green, P. J., 1998. Exact sampling from a continuous state space. Scandinavian Journal of Statistics 25, 483–502.

[12] Nishimura, K. and J. Stachurski, 2010. Perfect simulation of stationary equilibria. Journal of Economic Dynamics and Control, 34, 577–584.

[13] Propp, J. G., Wilson, D. B., 1996. Exact sampling with coupled Markov chains and applications to statistical mechanics. Random Structures and Algorithms 9, 223–252.