# A Real Options Perspective
# on the Future of the Euro[*]

by
Fernando Alvarez, University of Chicago
and
Avinash Dixit, Princeton University

This version: May 29, 2013

### Abstract

A break-up of the Eurozone is no longer regarded as implausible. This will be a costly and irreversible decision in conditions of continuing uncertainty; therefore it is amenable to analysis in the real options framework. We do so by solving as an $n-$dimensionaloptimal stopping problem with country-specific shocks and "convergence" of member economies. We compare a complete break-up with individual country departures. In calibrated solutions for a symmetric case we find a non-negligible but small option value. Furthermore, we find a new theoretical result on the non-monotonicity of abandonment threshold with respect to volatility.

**E-mail addresses of authors:**
    Alvarez: `f-alvarez1@uchicago.edu`
    Dixit: `dixitak@princeton.edu`

# 1  Introduction

The possibility that the Eurozone may break up, either completely, or partially through voluntary or forced exit of one or more member countries, is no longer regarded as impossible. It has shifted from being unthinkable to most people at the onset of the Great Recession to being considered unavoidable by many by 2011. While fears of break-up have receded after the official commitment by the European Central Bank on the irreversibility of the euro, the genie is out of the bottle. The up-front costs of such a break-up are recognized to be high, but the benefits to some countries of being able to pursue independent monetary policies better tailored to their distinct circumstances may outweigh the loss of benefits of membership of a large currency area. However, there is considerable uncertainty about future economic shocks and therefore about the balance of future flows of benefits. Therefore any break-up becomes a decision involving irreversible immediate costs and uncertain future benefits.

The theory of such decisions has been developed in considerable detail and has found many applications; expositions include Dixit and Pindyck (1994), Trigeorgis (1996), Stokey (2008), and Chevalier-Roignant and Trigeorgis (2012). The most important general insight is that most such decisions are not of a now-or-never nature; they include an option to wait for better information. Therefore they have been amenable to analysis using methods similar to those used for pricing financial options, and using that analogy they have been called real options. In this paper we explore these methods to model the break-up decisions for the euro.

The optimal timing of any break-up is an important and frequently discussed issue, but there is surprisingly little theoretical literature on it from the real option perspective. The only closely related paper we have been able to find is Strobel (2007), who modeled the decision of one country to join the European Monetary Union, but the break-up decisions involve different considerations. The work by Fuchs and Lippi (2006) address in a rich dynamic game set-up the issue of break-up. In their two country model, each country utilities flow are subject to random variation, and thus the break-up does include an option value. Yet given the complexity of the strategic interactions their model is parameterized in relatively simple way and not easily connected with plausible empirical magnitudes.[1]

The general problem is quite complex to analyze theoretically, and we have simplified it drastically to obtain some initial intuitions and results. Our main two simplifications are

---

[1]Because of the complexity of the game theoretical set up, Fuchs and Lippi (2006) consider iid shocks. Transitory shocks unrealistic for a quantitative evaluation or our application and, as we explain below, they also have important effects on the measurement of option values.

to use a reduced-form model to capture the behavior of the private sector and to consider the maximization problem for the collective of all country members. The model of the flow costs of a monetary union is in the Mundell-Fleming tradition. The flow cost from the lack of independent monetary policy is given by a quadratic loss function of the "nominal misalignment" of each country, which are assumed to followed an exogenous persistent process while the union is maintained. We interpret the nominal misalignment as the cumulative inflation differential that optimal country-specific monetary policy would have eliminated. Furthermore, we allow for a union-wide optimal monetary policy, which influences all countries alike, and hence ends up eliminating the union-wide average misalignment and leaving individual specific misalignments. On the other hand, we assume that membership to the union gives a flow benefit, independent of monetary policy, which are lost with a break-up. While it will be useful to attempt a better macroeconomic structural specification, our first efforts at that have proved intractable. We also make some restrictive assumptions such as symmetry of structure (of course not symmetry of realization of shocks), which allows us to manage a multidimensional optimal stopping problem. The assumption of the collective decision means that we can consider the option of either keeping the union or breaking up completely. We complement our analysis with an evaluation of the incentives of one country to leave the union. All these assumptions are explained in detail at the appropriate points below. Overall, while our model is simplistic about the intricacies of the costs and benefits of a multi-country currency union as well as of its dynamics, we think that the analysis of option value that we provide can be applied to existing models which give a more satisfactory representation of some of these features but which nevertheless ignore the option value of optimally timing a break-up.

We define the "now or never" strategy as one that equates the cost of break-up with the net present value of staying *forever* in the union. As stated above, the insight from the real option literature is that for an action that is irreversible and/or that involves a fixed cost, such as the break-up of the eurozone, the simple net present value calculations of the "now-or-never" case gives the wrong answer. By definition the "now or never" strategy does not incorporate the "option value" of waiting to see if conditions improve. Instead, using the optimal decision rule when the union is indifferent between breaking up and continuing the present value of staying forever in the union is strictly lower than the fixed cost of break-up. In other words, the average misalignment that the union tolerate is larger than what is prescribed in the nor or never case. Our main finding is that for the benchmark values of the parameters that we consider, the option value to wait is not large. More precisely, treating the decision of a break-up as a now-or-never choice leads to a non-negligible but small economic

loss, say on the order of a one time 4% annual GDP loss per country. Equivalently, the threshold for abandonment in the now-or-never decision, implies a typical misalignments in each country, interpreted as the average cumulated inflation differentials, which is about 5% smaller than the optimal threshold which takes into account the option value. We conclude that Eurozone will not be making a big mistake if it overlooks the option value and abandons the Euro as soon as the equivalent of the now-or-never threshold is reached.

Unsurprisingly, we find that in the absence of transfers among countries, the exit of a country with a large enough misalignment will be individually optimal even though the optimal aggregate union policy with transfers at this point will be not to break-up the union. The single country's optimum can be aligned with the aggregate optimum either if the cost of its individual exit is made disproportionately larger, or its benefit from staying in are disproportionally larger. Somewhat surprisingly we find that for a large set of parameters the magnitudes of these extra penalties (or bonuses to stay in) are relatively small.

Finally, we show that the optimal abandonment threshold can be either an increasing or a decreasing function of the volatility parameter $\sigma$, depending on whether the speed of mean reversion of all countries is smaller than the discount rate. Indeed for our preferred benchmark values we find the threshold to be decreasing on $\sigma$. At first, this seems to contradict the conventional wisdom about option values, namely that they are more important when there is greater uncertainty. However, the reason is that the abandonment threshold that would be employed if the only choice available were to abandon now or never is itself a decreasing function of $\sigma$. Interestingly, and intuitively, we show that the strength of the *wait* and see effect of $\sigma$ into the abandonment threshold depends on the speed of mean reversion. Nevertheless, the difference between the optimal threshold and the now-or-never threshold, which captures what we refer to as the *pure* option value of waiting, is an increasing function of $\sigma$, but as explained, this aspect may not be strong enough to overcome the effect of $\sigma$ on the now-or-never threshold. This generalizes the known results on the option value of waiting, clarifies exactly what the option value of waiting should mean, and provide a intuitively and simple characterization for the non-monotonicity of thresholds, with potentially applications in many other problems.

## 2   The model

The currency union member countries are labeled $i = 1, 2, \ldots n$. Each country $i$ has a state variable $X_i$, which measures the gap from the ideal monetary target for country $i$ in the absence of any monetary policy. In our implementation we will use the departure of the

country's real exchange rate from its long run average PPP level. Each state follows the dynamics

$$dX_i = -\mu_i \; X_i \; dt + \sigma_i \; dw_i + \sigma_c \, dw_c \,, \tag{1}$$

where the $w_i$ are standard Brownian motions. The $w_i$ for $i = 1, 2, \ldots n$ are country-specific shocks and $w_c$ a common shock, with

$$\mathrm{E}[dw_i] = 0, \; \mathrm{E}[dw_i^2] = dt \text{ for all } i, \text{ and } \mathrm{E}[dw_i \, dw_j] = 0 \text{ for all } i \neq j. \tag{2}$$

The parameter $\mu_i > 0$ is the rate of convergence of $X_i$ to its unconditional mean, normalized to zero. The case of $\mu_i = 0$ corresponds to a random walk. The parameter $\sigma_i^2$ gives the variance of the (innovation) into the country specific component and $\sigma_c^2$ the variance of the innovation common to all countries.

We denote the (effect of) monetary policy in country $i$ by $Z_i$. If each country $i$ can set its own monetary policy independently, its central bank could use it to set the country's deviation $x_i \equiv X_i - Z_i = 0$, the ideal value at all times. We assume that if the currency union is broken, each country will follow that policy. In the European union, the common central bank must have the same policy for all, say $Z = Z_1 = \cdots = Z_n$. With a common policy the countries' deviations are $x_i \equiv X_i - Z$.

Country $i$'s flow benefit from belonging to the Eurozone is

$$u_i(x_i) = \alpha_i - \tfrac{1}{2} \, \beta_i \, x_i^2 \,, \quad \text{with } \alpha_i \geq 0 \ \text{ and } \ \beta_i \geq 0 \,. \tag{3}$$

The benefit of membership under ideal conditions ($x_i = 0$; no exchange rate misalignment) for country $i$ is $\alpha_i$. Convexity of the loss due to misalignment is intuitive; then the quadratic is just a simple way to capture it.

We assume that abandonment of the common currency[2] entails a lump-sum cost $\Phi$. After paying this cost all countries switch to having separate currencies: they lose the $\alpha$'s, but assuming ideal conduct of monetary policy by their separate central banks, they get $u_i = 0$ for ever after. If this happens, it is difficult to imagine successful launch of another such project in any foreseeable horizon, so we assume this step to be essentially irreversible. We concentrate most or our analysis in the case where, given the assumptions about the collective problem we will make below, the union will either continue or break-up completely. In Section 6 we consider a single country's decision to exit, and compare this with the collective optimum.

---

[2]This has been recently referred to as *eurogeddon,* a term that made the short-list for the Oxford English Dictionary's best new word for 2012

We assume that the behavior of the private sector in each of the $n$ countries, captured by the parameters in the law of motions of each $X_i$, as well as the effect on the welfare of each country, captured by the parameters in each $u_i$, are independent of the process for the common monetary policy $\{Z\}$, and also of the prospects and timing of the abandonment of the union. Note that our formulation abstracts from the role of any further intertemporal links in the private sector decisions constraining the decisions of the collective, as well as from any reaction of the private sector to future monetary policies, including the abandonment itself.

We interpret the problem of the collective as follows. We assume that transfers are available among members countries – therefore we use the sum of the utilities as objective function – and that they can commit to follow the stopping rule decided by the collective for the abandoning of the union. The closest theoretical benchmark to the collective problem before abandonment will be Werning and Farhi's (2012) fiscal unions among $n$ countries in a sticky price or sticky wage set up. The closest analogy with the Eurozone would be a fiscal pact between the country members and the joint issuance of eurobonds. We briefly explore in other sections the implications for lack of commitment and characterize the incentives of a single country to deviate.

Given the assumed commitment and availability of transfers, the decision of the collective can be cast as a maximization problem: to choose the process for the optimal common monetary policy while the union is in place, $\{Z(t)\}$, as well as the state-dependent stopping rule ("stopping time") to abandon the union, $\{\tau\}$. The maximized value for the collective is

$$
\mathcal{V}\left(X_1, ...., X_n\right) =
$$
$$
\sup_{\{\tau \geq 0, Z\}} \mathrm{E}\left[ \int_0^\tau \sum_{i=1}^n u_i\left(X_i(t) - Z(t)\right) e^{-rt}\, dt - e^{-r\tau}\, \Phi \mid X_i(0) = X_i,\, i = 1, ..., n \right]
$$

subject to the law of motion of each state in equation (1) and the initial condition $X_1, ...., X_n$. This is a $n$ dimensional optimal stopping problem, whose solution consist on finding the subset of $\mathbb{R}^n$ that contains the state of the collective for which break up of the union is optimal. This is a difficult problem to analyze, which we simplify by imposing the following assumption:

**Symmetry Assumption**: $\sigma = \sigma_i$, $\mu = \mu_i$, and $\beta = \beta_i$ for all $i = 1, ..., n$.

Of course symmetry is not true in the Eurozone; the member countries have very different sizes, labor market institutions, product market regulation, etc. Symmetry is the price we pay to obtain a solution for the collective that we can characterize. In this context we can think of the zone in terms of subgroups of similar-sized and perhaps similarly situated countries:

(1) Germany, (2) France and Belgium, (3) Italy, (4) Spain, Portugal and Ireland, and (5) the rest. Therefore in the numerical calculations that follow we have chosen $n = 5$. We will also select values of $\mu, \sigma$ and $\beta$ as to be suitable for a "typical" country.

Now we turn to the analysis of the collective problem under symmetry. This enables us to write the flow utility for the union as $U$, the sum of the $u_i$, as

$$U = n\,\alpha - \tfrac{1}{2}\,\beta\,Y\,. \tag{4}$$

where

$$Y = \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} (X_i - Z)^2\,. \tag{5}$$

Since $Z$ is chosen at each instant with no effect on the law of motion of the $X's$, it must maximize $U$, or equivalently minimize $Y$. Therefore

$$Z = \frac{1}{n}\sum_{i=1}^{n} X_i = \arg\min_{z}\sum_{i=1}^{n} (X_i - z)^2\,. \tag{6}$$

Then

$$
\begin{aligned}
dx_i &= dX_i - dZ \\
&= -\mu\,(X_i - Z)\,dt + \sigma\,dw_i - \frac{\sigma}{n}\sum_{j=1}^{n} dw_j
\end{aligned}
$$

Two conclusions from the optimal union policy are worth emphasizing. First, when the common policy is set optimally, the common shock $dw_c$ cancels out. Second, under the optimal policy, $\sum_i x_i = 0$ at all times. This is reminiscent of Werning and Farhi's (2012) optimal policy where the weighted average of wedges across countries is always set to zero under the optimal policy, or to Gali and Monacelli's (2008) result for output gap's across countries, both obtained in fully articulated models.

The resulting minimized $Y$ serves as the single one dimensional state variable for the collective's problem. The reduction to one state, albeit one with a non-linear law of motion, is what allows us to solve this problem. In Appendix A we show that it follows the diffusion process

$$dY = \left[(n-1)\,\sigma^2 - 2\,\mu\,Y\right]\,dt + 2\,\sigma\,Y^{1/2}\,dW\,, \tag{7}$$

where $W$ is a standard Wiener process. Note the term $(n-1)$ in the drift. In Appendix B we describe how the problem changes if there were no union-wide policy, i.e. if $Z = 0$, the result is quite intuitive: without a union-wide policy the problem is equivalent to one with the law of motion of one more country but with the same sum of the $\alpha$'s.

The optimal collective policy is to choose a state-dependent stopping rule, i.e. a stopping time $\tau$, to maximize

$$V(Y) = \min_{\tau} \text{E} \left[ \int_0^{\tau} U(Y(t)) \, e^{-rt} \, dt - e^{-r\tau} \, \Phi \; \bigg| \; Y(0) = Y \right] . \tag{8}$$

The solution follows the standard process of dynamic programming or the theory of real options, for example Karlin and Taylor (1981) chapter 15 or, Dixit and Pindyck (1994). There is a threshold $\overline{Y}$ such that no action is taken while $Y < \overline{Y}$, but the euro is abandoned when $Y$ reaches $\overline{Y}$ (or is abandoned immediately if $Y > \overline{Y}$ at $t = 0$).[3] In the range of inaction, the function $V$ satisfies the (Bellman) ordinary differential equation

$$\tfrac{1}{2} \, (2 \, \sigma \, Y^{1/2})^2 \, V''(Y) + [\, n \, \sigma^2 - 2 \, \mu \, Y \,] \, V'(Y) - r \, V(Y) + [\, n \, \alpha - \tfrac{1}{2} \beta \, Y \,] = 0 \,,$$

or

$$2 \, \sigma^2 \, Y \, V''(Y) + [\, (n-1) \, \sigma^2 - 2 \, \mu \, Y \,] \, V'(Y) - r \, V(Y) + [\, n \, \alpha - \tfrac{1}{2} \beta \, Y \,] = 0 \,. \tag{9}$$

At the abandonment threshold $\overline{Y}$ the value matching and smooth pasting conditions are satisfied:

$$V(\overline{Y}) = -\Phi, \qquad V'(\overline{Y}) = 0 \,. \tag{10}$$

The use of the sum of the squares of deviation as state variable for a symmetric control problem follows the analysis in Alvarez and Lippi (2012) of price setting for multi-product firms.[4] The fact that the threshold $\overline{Y}$ is the sum of the squares of individual member country deviations from the ideal implies that abandonment is triggered when a few countries suffer misalignments of large magnitude or when a smaller number of countries suffer misalignments of smaller magnitude.[5] However, the size of misalignments for few countries has to be less than disproportionately large relative to their number, because of the squaring. To illustrate this, let $\overline{x}_k$ denote the deviation that would trigger abandonment when each of $k$ of the $n$ member countries reaches it while the remaining $(n-k)$ stay at the ideal. Then

$$\overline{Y} = k \, (\, \overline{x}_k \,)^2 \,,$$

---

[3] Alvarez and Lippi (2012) used a verification argument to formally establish that this is indeed the form of the inaction of control sets. A straightforward modification of the argument can be used to establish the same property for the problem at hand.

[4] There are several formal differences between the problems. First, the process for the deviations $x_i$ in this paper features mean reversion. Second, the objective function has a positive constant, which change some of the comparative statics, such as the sensitivity to $r$ of the threshold. Third, this is not a recurrent problem, the problem ends after the break-up of the union. Finally, a small difference is the presence of the common shock and the union-wide policy $Z$.

[5] If some misalignments are positive, others must be negative, because the $x_i$ always sum to zero.

so for example

$$( \overline{x}_1 )^2 = 4 \ ( \overline{x}_4 )^2 \quad \text{or} \quad \overline{x}_1 = 2 \ \overline{x}_4 \, .$$

Although symmetry is a serious restriction in the above analysis and calculations, the outcome suggests a conjecture about more general cases. If countries differ in their parameters $\beta_i$, $\mu_i$ and $\sigma_i$, we conjecture that the region of inaction is ellipsoidal. Its axes along the individual country dimensions $i$ (which tell us how large a misalignment in this country by itself will trigger abandonment) will depend on these parameters. A country with higher $\mu_i$ will have a larger axis, i.e. a higher threshold justifying abandonment, because the misalignment is more likely to get corrected faster over time. A country with a higher $\sigma_i$ will have a larger option value of waiting, but as we will see in Section 4.1 below, this may or may not translate into a higher option-inclusive threshold. A country with lower $\beta_i$ will have a higher threshold, because it would take a larger misalignment for the flow benefit of this country to becomes sufficiently negative. Of course all these are only conjectures awaiting proof or disproof. The $\alpha_i$ are added up in the constant term, so in the collective decision the size of the ellipsoid will depend on their sum, but the axis lengths will not depend on the individual countries' $\alpha_i$.

We finish this section with a comment on the special case of $n = 2$ countries where, due to the common union policy $Z$, the flow objective function becomes linear in $Y = (X_1 - X_2)^2$, as derived in Appendix C. This simple characterization allows to examine the role of heterogeneity of $\alpha$'s, $\sigma$'s and $\beta$'s across the two countries on the value of $\overline{Y}$. First, obviously the problem can be written in terms of the sum of the $\alpha's$. Second, and more subtly, the problem for the collective with heterogeneity can be written as the one with homogeneity but replacing the mean of the $\sigma^2$'s instead of common value and the *harmonic* mean of the $\beta$'s instead of the common value, as shown in Appendix C. Thus dispersion of $\alpha's$ and $\sigma^2$'s is immaterial for $\overline{Y}$. Adding dispersion in $\beta$'s while keeping their sum constant decreases $\overline{Y}$, since we show below that the threshold is decreasing in the common value of $\beta$. This effect comes from the fact that $Z$ is tailored to respond more to the values of the countries with higher value of $\beta$'s. We conjecture that a similar effect will survive to the general case of $n \geq 2$.

# 3 Some illustrative numerical solutions

As in Alvarez and Lippi (2012), the equation can be solved by assuming a power series expansion

$$V(Y) = \sum_{m=0}^{\infty} c_m \, Y^m \tag{11}$$

and substituting into the equation (9). The details are in the Appendix D. Starting with an arbitrary $c_0$, the remaining coefficients can be calculated recursively. The $c_0$ and $\overline{Y}$ are then found using the value matching and smooth pasting conditions.

The equation can also be solved numerically by converting it into a finite difference equation. We tried both methods; fortunately they yield the same outcomes within small numerical errors.

We begin with solutions for a few values of the parameters that are artificial but serve to give some general intuitions and lead to some analytical work; then we turn to solutions for parameter values drawn from empirical literature.

Let $\alpha = 1$ and $\beta = 2$. These are normalizations amounting to choice of the units in which $V$ and $x$ (or equivalently time) are measured; thus each country's flow benefit from being in the Eurozone with an ideally aligned real exchange rate is taken to equal 1 ($u = 1$ when $x = 0$), and the departure from the ideal real exchange rate that annihilates this gain is also set equal to 1 ($u = 0$ when $x = 1$). We take the interest rate to be 5% per year, so $r = 0.05$. The lump sum cost of abandonment is set at $\Phi = 100$. Therefore if there were no convergence and no uncertainty (i.e. if $\mu = 0$ and $\sigma = 0$), the Eurozone would be abandoned when $U = -5$, that is, if each country's flow benefit fell to $-1$, so it was losing relative to being outside as much as it would gain under ideal conditions. That corresponds to the threshold $\overline{Y} = 10$. This should be taken as a basis for comparison of the results when $\mu$ and $\sigma$ differ from zero. Table 1 shows the abandonment threshold for various values of $\mu$ and $\sigma$.

Table 1: Abandonment threshold $\overline{Y}$

|  | | $\sigma$ | |
| --- | --- | --- | --- |
| $\mu$ | 0.0 | 0.2 | 0.3 |
| 0.0 | 10.00 | 12.85 | 13.58 |
| 0.0125 | 15.00 | 15.00 | 15.00 |
| 0.0250 | 20.00 | 18.57 | 17.32 |

Figure 1 illustrates the graph of the value function for $\mu = 0.025$ and $\sigma = 0.3$. Note that the rising portion to the right is only a formal continuation of the mathematical solution; it is not economically relevant as the abandonment threshold is at the point where the function reaches its minimum.
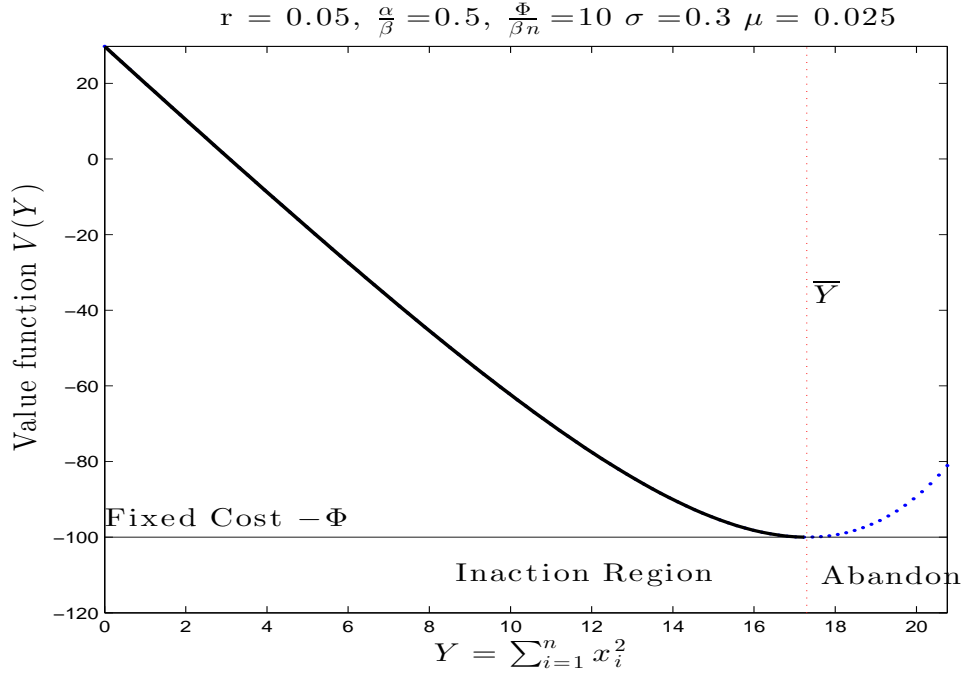


Figure 1: Graph of $V(Y)$ for illustrative parameter values.

# 4 Analysis of optimal threshold $\overline{Y}$

In this section we analyze the behavior of the threshold $\overline{Y}$ as functions of the parameters of the problem.

## 4.1 Comparison with now-or-never choice

In Table 1, for $\mu = 0$ the threshold $\overline{Y}$ increases as $\sigma$ increases; this conforms with the usual intuition that the option to wait is more valuable when there is more uncertainty. However, when $\mu = 0.0125$ the threshold stays the same, and when $\mu = 0.025$ the threshold decreases as $\sigma$ increases. This is surprising at first sight as it goes against that intuition and the analysis of most real option problem –see for example the cases analyzed in Dixit and Pindyck (1994) and Stokey (2008). However, the threshold in this table is the overall result of two distinct

effects of uncertainty: one on the threshold that would be optimal if the only choice available were either to abandon right away or to stay in for ever, and the other the additional effect of the availability of the option to wait and postpone the decision. To isolate these two effects, we compute the value of $Y(0)$ that will leave the collective indifferent between abandoning now or preserving the union forever, and refer to it as $\widehat{Y}$. While the thought experiment behind $\widehat{Y}$ is forward looking, by definition it does not include the option value of waiting. An expression for $\widehat{Y}$ is easy to calculate; see Appendix E for the details. The result is

$$\widehat{Y} = \max\left\{\frac{2\,(2\mu + r)}{r\,\beta}\,[\,r\,\Phi + n\,\alpha\,] - \frac{(n-1)\,\sigma^2}{r}\,,\ 0\right\}. \tag{12}$$

This is a decreasing function of $\sigma$, as long as $\widehat{Y} > 0$. That makes intuitive sense: if $\sigma$ is larger, there is a bigger probability of drifting into a range of high $Y$; the flow losses from this are large because of the convexity of the cost function. In the now-or-never decision, if one does not abandon now, one is stuck with this risk for ever after. Therefore it is better to abandon at a lower $Y(0)$. For $\sigma$ large enough there is no value of $Y$ for which the collective is indifferent, and hence $\widehat{Y} = 0$. Likewise it is an increasing function of $\mu$, since for higher values the collective misalignments are expected to decline exponentially at a higher rate, making it less desirable to pay the up-front cost of abandonment.

Table 2 shows the comparisons between the overall or option-inclusive threshold $\overline{Y}$ and the now-or-never threshold $\widehat{Y}$. In each cell the option-inclusive threshold is listed first, from that is subtracted the now-or-never threshold, and the pure option value effect is the result, shown to the right of the $=$ sign. All calculations are for the same parameters as used above: $n = 5$, $\alpha = 1$, $\beta = 2$, $r = 0.05$, $\Phi = 100$.

Table 2: Option-inclusive versus now-or-never thresholds: $\overline{Y} - \widehat{Y}$

| | $\sigma$ | | |
|---|---|---|---|
| $\mu$ | 0.0 | 0.2 | 0.3 |
| 0.00 | $10.00 - 10.00 = \mathbf{0.00}$ | $12.85 - 6.80\ \ = \mathbf{6.05}$ | $13.58 - 2.80\ \ = \mathbf{10.78}$ |
| 0.0125 | $15.00 - 15.00 = \mathbf{0.00}$ | $15.00 - 11.80 = \mathbf{3.20}$ | $15.00 - 7.80\ \ = \mathbf{7.20}$ |
| 0.0250 | $20.00 - 20.00 = \mathbf{0.00}$ | $18.57 - 16.80 = \mathbf{1.77}$ | $17.32 - 12.80 = \mathbf{4.52}$ |

For any given $\mu$, the pure option effect is increasing in $\sigma$, confirming the usual intuition. For $\mu = 0$, the pure option effect increases so rapidly with $\sigma$ that the total threshold also increases. This also makes sense: if $\mu$ is small, convergence is not going to be of much help to

reduce the flow costs of misalignment over time. Therefore there is high value of waiting to see if a random fluctuation moves the economy in the right direction, that is, a high option value. For $\mu > 0.0125$, the option effect is not strong enough to offset the negative effect of $\sigma$ on the now-or-never threshold, so the total threshold decreases as $\sigma$ increases. When $\mu = 0.0125$ the total threshold $\overline{Y}$ is independent of $\sigma$. We follow this up with some analytical work in Section 4.2, and find that such is indeed the case when $r = (n-1)\,\mu$. (In the above work, we have $r = 0.05$ and $n = 5$, so the equality holds when $\mu = 0.0125$.)

## 4.2   Comparative Static of $\overline{Y}$

In this section we explore how the threshold $\overline{Y}$ depends on the six parameters of the collective problem $n$, $\alpha$, $\beta$, $\Phi$, $r$, $\mu$ and $\sigma^2$. In particular, i) we show that the six parameters determined $\overline{Y}$ can be combined into four, ii) we give a closed form expression for $\overline{Y}$ for small $r/\sigma^2$ and $\mu/\sigma^2$, iii) we characterize the surprising comparative static of $\overline{Y}$ with respect $\sigma^2$, and finally iv) we characterize the remaining intuitive comparative static of $\overline{Y}$ with respect to $\alpha$, $\beta$, $\Phi$, $\mu$ and $r$.

First we develop three homogeneity properties that imply that only four parameters matter for $\overline{Y}$. Inspecting the objective function it is immediate to see that it is homogeneous of degree one in $(\alpha, \beta, \Phi)$, and hence the optimal threshold $\overline{Y}$ is homogenous of degree zero (and for the same reasons so is the now-or-never threshold $\widehat{Y}$). Thus $\overline{Y}$ can be written as a function of the ratios $(\alpha/\beta, \Phi/\beta)$. Second, fixing $r$, $\mu$, $\sigma$, $\beta$ the threshold depends only on $\Phi + \alpha\, n/r$, the sum of the fixed cost and the present value of the flow benefit of belonging to the union which is lost at abandonment. This is quite intuitive since breaking up the union imposes both costs, and hence their composition is unimportant. Formally for each path we can rewrite the objective function in equation (8) as:

$$\int_0^\tau \left[ n\alpha - \frac{\beta}{2} Y(t) \right] e^{-rt}\, dt - e^{-r\tau}\, \Phi = \frac{n\,\alpha}{r} - e^{-r\tau} \left[ \frac{n\,\alpha}{r} + \Phi \right] - \int_0^\tau e^{-rt} \frac{\beta}{2} Y(t)\, dt$$

and hence $\overline{Y}$ (and for the same reason $\widehat{Y}$) depend only $\Phi + n\,\alpha/r$. Third, the threshold is independent on the units at which time is measured, so if $r, \mu, \sigma^2, \alpha, \beta$ are multiplied by a positive constant $\lambda$ then $\overline{Y}$ remains the same. Setting $\lambda = 1/r$ we can write the threshold ss a function:

$$\overline{Y} = \varphi \left( n\,, \; \frac{r\Phi}{\beta} + \frac{\alpha}{\beta} n\,, \; \frac{\mu}{r}\,, \; \frac{\sigma^2}{r} \right) . \tag{13}$$

Second, for small values of $\mu/\sigma^2$ and $r/\sigma^2$, we have the following analytical approximation for the value of $\overline{Y}$:

$$\bar{Y} \approx 2\frac{n+1}{n-1} \left( \frac{n\,\alpha + r\,\Phi}{\beta} \right) + 16 \frac{n+1}{(n+3)(n-1)} \left( \frac{\alpha}{\beta} \right)^2 \left[ \frac{(n-1)\mu - r}{\sigma^2} \right] . \tag{14}$$

13

This approximation is developed in Section G where we study the undiscounted problem (i.e. $r = 0$) and use it to evaluate a Taylor expansion of the general solution of $\overline{Y}$ around $r = 0$ and $\mu = 0$. The approximation in equation (14) depends on four parameters as indicated in equation (13), but being an approximation it satisfies the exact form of equation (13) only for $r/\sigma^2 = \mu\sigma^2 = 0$. Moreover the approximation in equation (14) confirms that the pattern displayed in Table 1 for particular the parameter values of that table: it shows that $\overline{Y}$ is increasing in $\sigma$ when $(n-1)\mu < r$, decreasing in $\sigma$ if $(n-1)\mu > r$, and independent of $\sigma$ when $\mu(n-1) = r$.

Third, we explore in the general case the behavior of $\overline{Y}$ with respect to $\sigma^2$. Consistently with the approximation for small values of $r/\sigma^2$ and $\mu/\sigma^2$ and the numerical results of Table 1 we show that whenever $r = (n-1)\mu$ the partial derivative of $\overline{Y}$ with respect to $\sigma$ is zero. The proof of this result can be found in Appendix I, and its logic is as follows. The solution to the collective decision problem is given by the value matching and smooth pasting conditions (10). They implicitly define the threshold $\overline{Y}$ and the coefficient $c_0$ in the power series solution (11) to the differential equation (9) as a function of all parameters. We obtain the desired result by totally differentiating these equations and solving for the changes in $\overline{Y}$ and $c_0$ with respect to $\sigma^2$.

An intuitive explanation of the result for the derivative of $\overline{Y}$ with respect to $\sigma^2$ can be obtained by considering two extreme cases. The first is the standard case of $\mu = 0$, where the threshold is increasing in volatility[6]. The explanation in this case is that if one were to keep the threshold *constant* in the face of higher volatility $\sigma$, the value function will increase, at least for values $Y$ lower but close to $\overline{Y}$ thus implying a higher value of the optimal threshold. To see why the value function increases with $\sigma$ consider the case where $Y = \overline{Y}$: note that if after a shock $Y > \overline{Y}$ the union is abandoned and the payoff is the same as in the case with lower volatility, but if after a shock $Y < \overline{Y}$ then with higher volatility the process $Y$ lands at lower values which correspond to higher flow payments for the union, thus increasing the level of the value function. A second extreme is the one of $\mu \to \infty$ which in discrete time corresponds to an i.i.d. process for $x_i$ and thus for $Y$. In this case, the current position of $Y$ has no effect on the value of $Y$ in the subsequent period and thus higher volatility only decreases the value function since the flow payment is convex in the deviations $x_i$ (see Appendix J for a formal analysis of the iid case). Thus in general there are two effects of the volatility $\sigma$ into $\overline{Y}$. The first effect is the value of waiting to see if things improve, which tends to make $\overline{Y}$ increasing in $\sigma^2$. The second effect comes from the assumption that the cost

---

[6]See Dixit (1991) for a proof in a closely related problem.

of deviation are convex (i.e. $\beta > 0$), and thus the cost of continuing for ever increases when volatility increases. Therefore the now-or-never threshold decreases as volatility increases. This effect tends to make $\overline{Y}$ decreasing in $\sigma^2$. When $\mu$ is large, the high mean reversion reduces the effect of waiting for the shock to reverse. Therefore the option value effect on the optimal threshold becomes less important. Putting these together, as $\mu$ increases the threshold $\overline{Y}$ goes from being an increasing to a decreasing function of the volatility $\sigma^2$. The precise dividing line, namely $\mu = r/(n-1)$, must of course be calculated out and cannot be guessed by intuition alone.

Forth, we show five comparative static results of the optimal threshold $\overline{Y}$ with respect to its determinants:

1. $\overline{Y}$ is (weakly) increasing in $\alpha$,

2. $\overline{Y}$ is (weakly) decreasing in $\beta$,

3. $\overline{Y}$ is (weakly) increasing in $\Phi$,

4. $\overline{Y}$ is (weakly) increasing in $\mu$, and

5. $\overline{Y}$ is (weakly) increasing in $r$ for $\alpha$ small enough.

The proof of these results are in Appendix F. Note that these results hold for the approximation in equation (14). The first four results are straightforward. Since $\alpha$ is the constant flow benefit of staying in the union and $\Phi$ the fixed cost of abandoning the union, it is intuitive that larger misalignments are tolerated when these parameters are larger, as stated in 1 and 3. Since $\beta$ measures the cost of a given misalignment, it is intuitive that a smaller one is tolerated, as stated in 2. Since higher $\mu$ implies that costly mis-alignments self-correct at a faster rate, then it is intuitive that higher ones are tolerated, as stated in 4. Finally, when $\alpha$ is small the expected discounted future flow benefits close to the optimal threshold are negative, and thus a higher discount rate makes them less important.

# 5    A calibrated example

Now we switch from parameter values that are useful for illustrating general conceptual points to ones that are guided by both empirical research and stylized versions of existing models that can be mapped into our simple framework. First we give a brief discussion of the motivation for our choice of benchmark parameters values (for more details see Appendix K,

where we review the variables which we used to measure the extent of the option value, and finally we present several graphs with measures of the option value).

We interpret $x$ to be a misalignment of real exchange rates that can – and should – be "corrected" by an appropriate monetary policy. We measure $x$ as deviation from PPP across countries, for which there is a large empirical literature establishing that such deviations are large and very persistent. The one year standard deviation of changes in real exchange rates is between 6% and 10% for developed countries. The half-life of relative PPP deviations is at least between 3-5 years across countries, but the estimates are very imprecise, so much larger half lives are hard to distinguish statistically. Using the formula for variance of yearly changes, namely $\sigma^2 \left(1 - e^{-\mu}\right)/\mu$, and for the half-life, $\log(2)/\mu$, we set our baseline parameters to $\sigma = 0.08$ and $\mu = 0.1$.

We measure the parameters $\Phi/n$ and $\alpha$ as fractions of country GDP. For the flow benefit $\alpha$ we include two types of considerations: the gains from the increased trade as well as the reduction in transaction cost. Based on this consideration we set $\alpha = 0.02$, with about a quarter of this is due to transaction-cost reductions and the rest comes from the increase in trade.

For the value of $\Phi$ we rely on the recent experience of countries banking/currency/debt crises. We identify $\Phi$ as the cost resulting from re-introducing a new currency, and most importantly dealing with the likely defaults and disruption into the financial and payment system that this may cause. We set $\Phi/n = 0.2$.

Finally we discuss the parameter $\beta$, the sensitivity to square deviation of $x_i$. This is an important parameter which requires a fully specified model to have a clear interpretation. Specifying a dynamic model for the determination of $x_i$ which will be helpful for the measurement of the misalignment and its welfare consequences – i.e. $\beta$ – and be able to solve for optimal threshold goes beyond the scope of this paper. Instead we motivate our choice of $\beta$ with a simple static model, with tradable and non-tradable goods, and factors freely moving between the sectors. As often used in the sticky price literature, we consider the deviation from PPP as if it were an equivalent tax or "wedge", so we compute the equivalent variation in consumption that will make the country indifferent between the undistorted allocation, which corresponds to a value of $x = 0$ and the one with relative price $e^x$. In the case of two symmetric goods (i.e. domestic and foreign) with constant elasticity of substitution $\eta$ and a share of tradable goods $\epsilon$, this model gives $\beta = \epsilon\eta$. Thus, using a tradable share $\epsilon = 0.3$ and an elasticity $\eta = 6$ we obtain $\beta \approx 2$. An alternative will be to use a version of the fiscal union analyzed by Werning and Farhi (2012) based on the static model by Obstfeld and Rogoff

(1995) and converting the labor labor wedge into our state. We leave the analysis of this exploration for future work.

As mentioned in Section 2 we use $n = 5$ for our calculations, dividing the Eurozone in 5 regions of approximately the same size, as required by our model. We use the transformation $(Y/n)^{1/2}$ to express thresholds $\overline{Y}$ and $\widehat{Y}$ in units of deviation of $x$ for a typical country. Since $Y = \sum_i x_i^2$ and each $x_i$ is country's $i$ real exchange rate, i.e. its deviation from PPP, then the transformation $(Y/n)^{1/2}$ expresses $Y$ into units of a typical country deviation from PPP, which is exact if all the deviations are of equal absolute value across countries. Hence the units of this transformation of $Y$ can be thought as cumulative inflation differentials for a typical country. Alternatively, if all the misalignment were to be as concentrated in one country as possible -recall that all deviations must add to zero- so we can express it in units of $x$ as $(Y(n-1)/n)^{1/2}$, which relative to our previous measure is $(n-1)^{1/2}$ times larger, or twice as large for our choice of $n = 5$. As a reference, these values can be compared with the benchmark values for the standard deviation of either the yearly innovations in $x$, which is $\sigma = 0.08$, or the unconditional standard deviation, which equals $\sigma/\sqrt{2\,\mu} \approx 0.18$.

We normalize 1 to be the annual GDP of each of the $n$ symmetric countries if they were to abandon the union; therefore the GDP of the whole area is simply $n$. Thus, $\alpha = 0.02$ has the interpretation of an annual flow benefit of belonging to the union of 2% of the annual GDP for each country. $\Phi$ is the break-up cost for the whole area, so that $\Phi/n$ has the units of fixed cost as percentage of each country GDP. For instance, $\Phi = 0.20 \times 5 = 1$ for $n = 5$ means that the break-up cost is equivalent to a one-time reduction of 20% of each country GDP for a period of a year. Finally, if country's $i$ has a misalignment of $x_i$ during a year, then its welfare decreases by $\frac{1}{2}\beta\,x_i^2$ measured in country's $i$ equivalent annual GDP units. For instance, for $\beta = 2$ if a country has $x_i^2 = 0.05^2$ for a whole year, its welfare decreases by one quarter of one percent of annual GDP.

Finally, consider the gain from using the optimal policy instead of the now-or-never policy. Let $V_E(Y)$ denote the discounted present value of starting at $Y$ and continuing the union for ever after. An expression for this is derived in Appendix E in equation (A-9). The now-or-never threshold $\widehat{Y}$ is defined there by the equation

$$V_E\left(\widehat{Y}\right) = -\Phi\,.$$

If, starting at $\widehat{Y}$, instead of abandoning immediately, we follow the optimal (option-inclusive) policy with its higher threshold $\overline{Y}$, the value will be $V(\widehat{Y}) > -\Phi$. Therefore $\left(V(\widehat{Y}) + \Phi\right)/n$ measures the gain from optimal use of the option to wait: the difference between value of
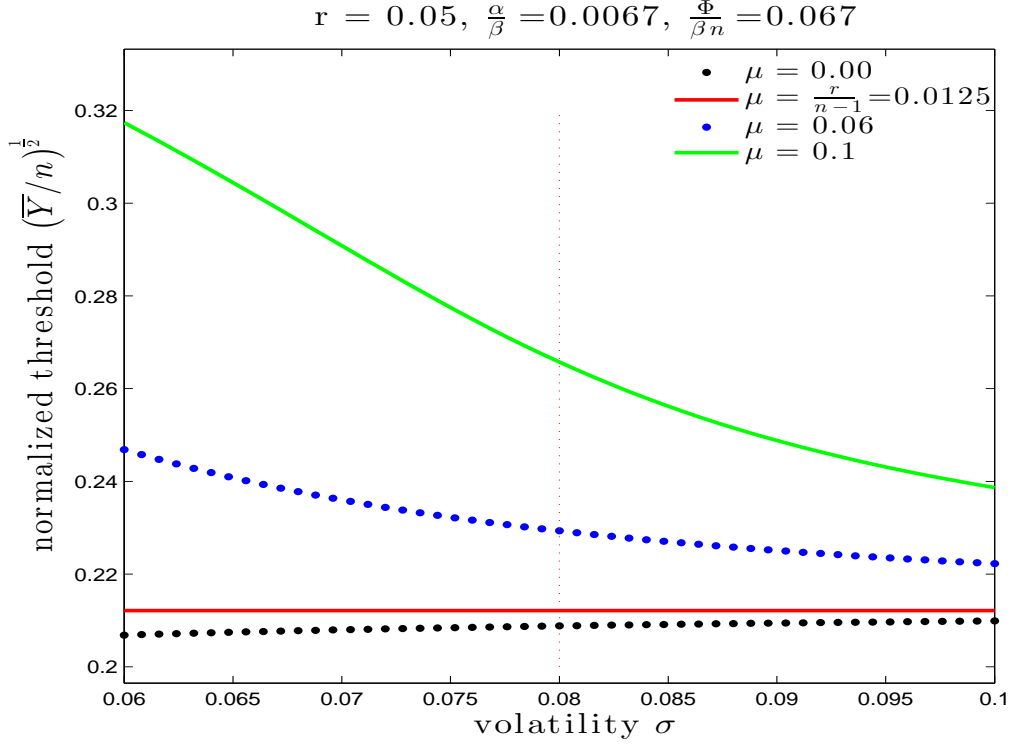
17

Figure 2: Normalized optimal threshold as function of $\sigma$ for selected $\mu$.

following the optimal policy relative to abandon the union at the point where that becomes optimal in the now-or-never case. Few remarks are in order. First, dividing the difference in the values by $n$ measures it in units of a typical country GDP. Second, this difference is evaluated at $Y = \widehat{Y}$, the value where the union will be indifferent between continuing or not if following a policy where abandonment is decided now or never. Finally, this measure of the option value has the interpretation of a once-and-for-all benefit as fraction of the typical country GDP, to be distinguished from a flow benefit to be enjoyed in perpetuity.

We start with Figure 2 which contains four lines: each of them corresponds to a different value of $\mu$ for which we display the level of the (normalized) optimal threshold $\overline{Y}$ as a function of $\sigma$. We use this figure to illustrate three points. The first one is the comparative static result for $\overline{Y}$ with respect to $\sigma$. As explained in Section 4.1 this figure illustrates that the behavior of $\overline{Y}$ as function of $\sigma$ depends on whether $(n-1)\mu \lesseqgtr r$. The highest line corresponds to our benchmark parameter values, for which $(n-1)\mu > r$ and hence $\overline{Y}$ is decreasing in $\sigma$. The lowest one correspond to $\mu = 0$, for which we obtain the standard result that $\overline{Y}$ increases with $\sigma$. The second point is the comparative static result with respect to the speed of mean reversion $\mu$. Note that for each $\sigma$ the value of the threshold is increasing in $\mu$. This

$$\text{r } = 0.05, \; \tfrac{\alpha}{\beta} = 0.01, \; \tfrac{\Phi}{\beta n} = 0.1, \; \mu = 0.1$$
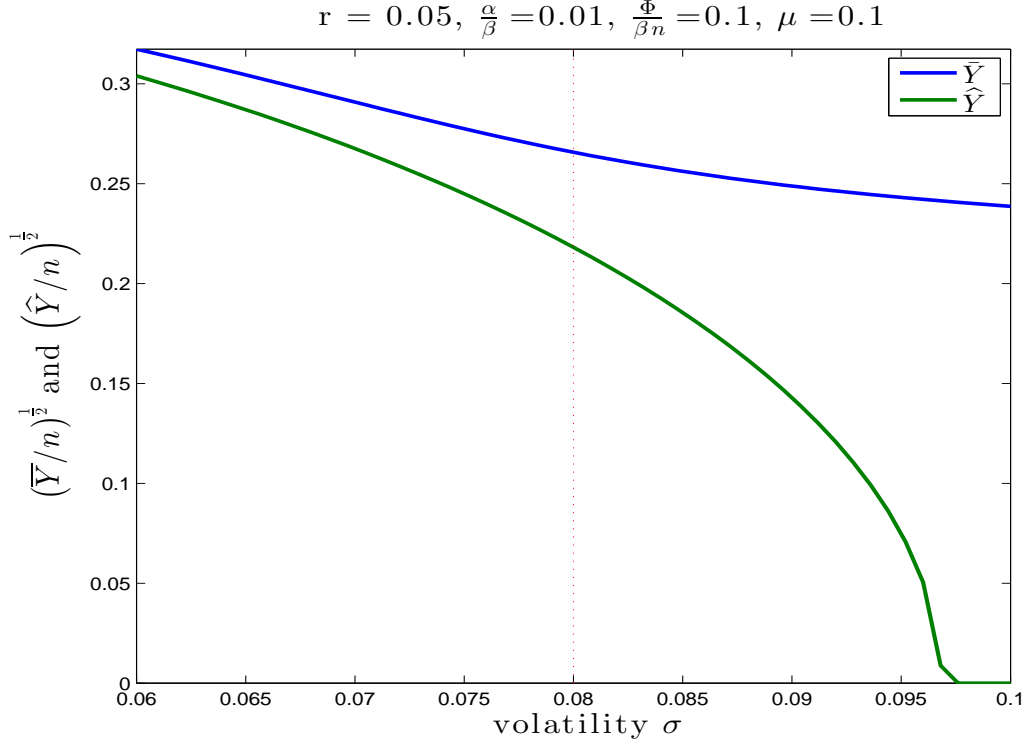
Figure 3: Normalized thresholds as function of $\sigma$ .

is to be both intuitive and in line with the approximate solution displayed in Section 4.1. If the misalignment will correct itself at a faster rate, the collective should be ready to tolerate higher misalignments before abandoning the union. The third point we make with this figure is about the size of the implied corrections on the misalignment in the event of a break-up of the union. For the benchmark values – indicated with a vertical dotted line – the size of the misalignment, which equals the correction at the time of abandoning the union, is large. If the misalignment were equal in the five regions the correction will be about 27%, and if it were concentrated in one region it will be about 60%, i.e. $0.27 \times 5^{1/2}$. As another reference, the level of normalized typical deviation $(\overline{Y}/n)^{1/2}$ is close to the value of the unconditional standard deviation of $x$. While these values are large they are well within the historical examples of observed exchange rate changes after severe banking and currency crises as the ones of Argentina in 2002, of Indonesia, Korea and Thailand in 1997, and of Russia in 1998.

We use Figure 3 and Figure 4 to assess the size of the option value. Figure 3 displays the normalized optimal and now-or-never thresholds as a function of $\sigma$ for our benchmark parameter values. As explained in Section 4.1, given our benchmark values, both thresholds are decreasing in $\sigma$, the value of $\overline{Y}$ is higher than the one for $\widehat{Y}$, and its difference is increasing in $\sigma$, as long as $\widehat{Y} > 0$. Note also that, for the largest values of $\sigma$ the value of $\widehat{Y}$ is zero.
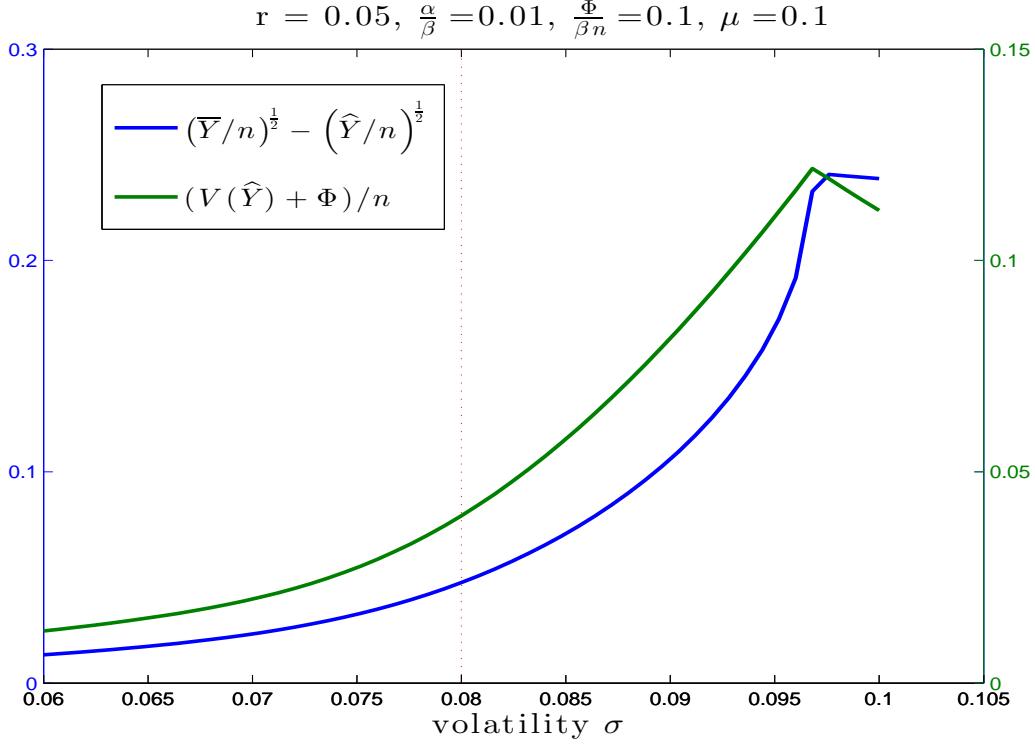
19

Figure 4: Two measures of the option value as a function of $\sigma$ .

Figure 4 displays two measures of the size of the option value. One is the difference between the two thresholds, which is about 6% if all countries have the same size deviation. The last segment of both lines behaves differently, since it corresponds to the cases where $\widehat{Y} = 0$. If the deviation were to be concentrated in one of the $n = 5$ groups (say Spain, Portugal and Ireland), the magnitude of the option value is about 13%, i.e. $0.06 \times 5^{1/2}$. This figure also displays the normalized difference in the value function achieved by following the optimal policy, which at the benchmark parameter values is about 4%.

As shown in Section 4.2 the thresholds $\overline{Y}$ and $\widehat{Y}$ depend only on the ratio $(n\,\alpha/r + \Phi)\,/\beta$, and the value function $V$ is homogeneous of degree one in $(n\,\alpha/r + \Phi\,,\,\beta)$. Figure 5 shows how variation on $\beta$ affects the levels of the normalized thresholds and Figure 6 shows how it affects the two measures of the option value. As expected, if misalignment is more costly, i.e. $\beta$ is higher, then the thresholds are smaller but the option value is more important. Even for the large range of $\beta$ in these figures our estimates of the $\overline{Y}$ and of the option value do not seem to change dramatically.

From these numbers it appears that the difference in the thresholds and values corresponding to the now-or-never decision and the truly optimal decision, while significant, is
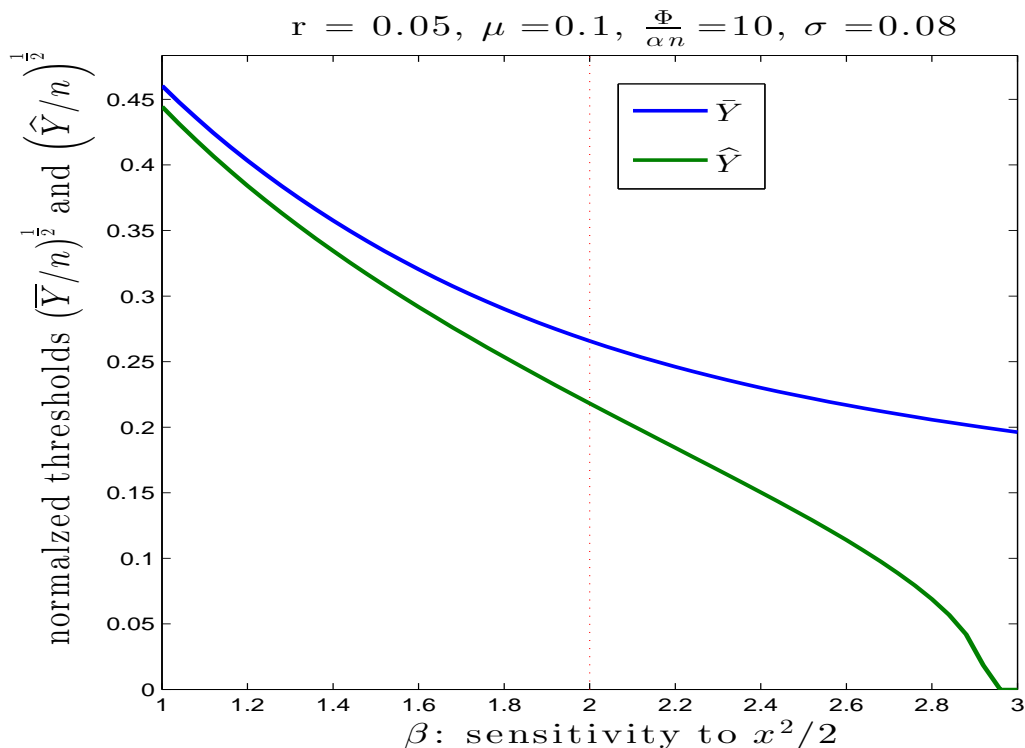
Figure 5: Normalized thresholds as function of $\beta$ .

not large. If we still believe in "convergence," we should use a larger $\mu$ and/or a smaller $\sigma$ than our benchmark numbers, and for these the pure option value will be even smaller. Moreover, we have assumed that individual countries' monetary policies can eliminate all deviation from PPP. This is unrealistic; there are quite large real exchange rate fluctuations in countries with independent monetary policy.[7] To summarize, we interpret the values found in this section as implying that the Eurozone will not be making a big mistake if it overlooks the option value and abandons the Euro as soon as the now-or-never threshold is reached.

# 6  One country's exit decision

The problem we have analyzed so far can be thought as the case where there are transfers across countries and commitment on the part of the collective, and hence the relevant criterion is to maximize the sum of the countries' utilities. In that interpretation, only the collective

---

[7]For instance, this will imply that variations of PPP across US states, which is a monetary union, have large negative welfare consequences. It will also imply that, following the optimal policy, countries that are not members of a monetary union should have a very stable real exchange rate, which is far from what is observed from most countries.
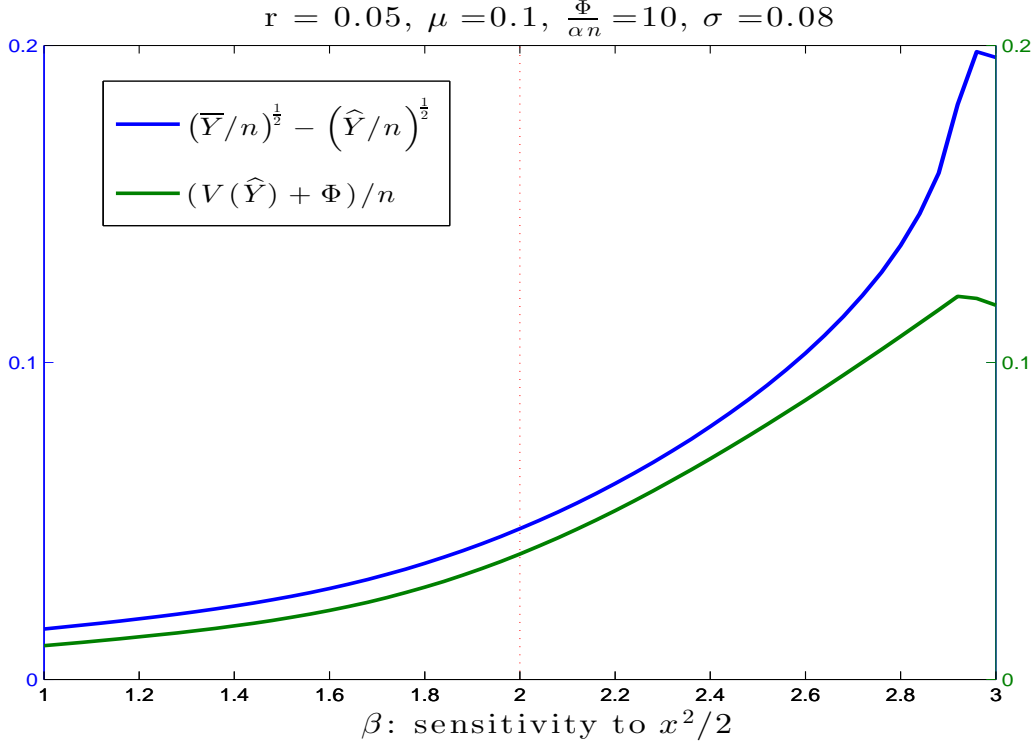
$$r = 0.05, \; \mu = 0.1, \; \frac{\Phi}{\alpha n} = 10, \; \sigma = 0.08$$

Legend:
- $\left(\overline{Y}/n\right)^{\frac{1}{2}} - \left(\widehat{Y}/n\right)^{\frac{1}{2}}$
- $(V(\widehat{Y}) + \Phi)/n$

$\beta$: sensitivity to $x^2/2$

Figure 6: Two measures of the option value as a function of $\beta$ .

state $Y$ matter. But suppose instead that no actual transfers are being made, and individual member countries can exit by paying lump sum cost, say $\phi$, so we have a non-cooperative dynamic game. This is even harder to analyze. In this section we shy away from that analysis and simply characterize the incentives of one individual country to abandon the union.

In particular we assume that the collective policy is to set $Z = (1/n) \sum_{i=1}^{n} X_i$ as long as $Y \leq \overline{Y}$ and abandoning the union the fist time at which when $Y$ reaches $\overline{Y}$, and consider the (expected discounted) utility for each of the country's members when this policy is followed. We inquire for which values of $\phi$ would an individual country find it optimal not to deviate from the policy followed by the collective. We find that $\phi \geq \Phi/n$, and that while for all the numerical examples the inequality is strict, the differences are not that large.

The flow benefit of country $i$ (now no actual transfers are being made) is

$$u(y_i) = \alpha - \tfrac{1}{2}\beta \, y_i \tag{15}$$

where

$$y_i = (X_i - Z)^2 \;, \quad \text{and} \;\; Y = \sum_{i=1}^{n} y_i \;. \tag{16}$$

Neglecting the individual country's subindex to simplify the notation, Appendix H shows that $y \leq (n-1)/n\,Y$ and that

$$dy = \left[ \sigma^2 \frac{n-1}{n} - 2\,\mu\,y \right] dt + 2\,\sigma\,\sqrt{y\frac{n-1}{n}}\; dw_y \text{ with } \mathrm{E}\,[\,dy\,dY\,] = 4\,\sigma^2\,y\;dt \quad (17)$$

and with $\mathrm{E}\,[\,dw_y\,dw\,] = [y/Y\,n/(n-1)]^{1/2}\,dt$. Define the present discounted value of a country belonging to the union when the country's state is $y$ and the collective state is $Y$ as

$$v(Y\,,\,y) = \mathrm{E}\left[\left. \int_0^\tau u(\,y(t)\,)\,e^{-rt}\,dt - e^{-r\tau}\,\frac{\Phi}{n}\; \right|\; Y(0) = Y\,,\,y(0) = y \right]$$

where $\tau$ is the first time that $Y$ reaches $\overline{Y}$. Note that for all $\{y_i\}$ and $Y$ satisfying (16) we have:

$$\sum_{i=1}^n v\,(Y\,,\,y_i) = V\,(Y)\;.$$

Given $\overline{Y}$ the function $v$ solves the following partial differential equation:

$$\begin{aligned}
rv(Y\,,\,y) &= \alpha - \frac{\beta}{2}y + \left[\sigma^2\,(n-1)/n - 2\,\mu\,y\right]\frac{\partial v(y,Y)}{\partial y} + \left[\sigma^2\,(n-1) - 2\,\mu\,Y\right]\frac{\partial v(y,Y)}{\partial Y}\\
&\quad + 2\,\sigma^2\,y\,\frac{\partial v(y,Y)}{\partial y^2} + 4\,\sigma^2\,y\,\frac{\partial v(y,Y)}{\partial y \partial Y} + 2\,\sigma^2\,Y\,\frac{\partial v(y,Y)}{\partial Y^2}
\end{aligned}$$

for all $0 \leq Y \leq \overline{Y}$ and $0 \leq y \leq \frac{n-1}{n}Y$. Since the union is dissolved when $Y = \overline{Y}$ we have the following boundary condition:

$$v\left(\overline{Y}\,,\,y\right) = -\Phi/n \text{ for all } 0 \leq y \leq \overline{Y}\;.$$

Given the symmetry of the different countries we have that when they all have the same misalignment, their values are the same, which gives the following relationship between the two functions: $n\,v\,(Y\,,\,Y/n) = V(Y)$ for all $0 \leq Y \leq \overline{Y}$.

Now consider a state such that the collective is indifferent between abandoning the euro or not, so $Y = \overline{Y}$ and where $n-1$ countries have identical small deviations, and one country has its largest possible deviation $y = Y\,(n-1)/n = \overline{Y}\,(n-1)/n$.

Indeed $\min_{y,Y} v\,(Y\,,\,y) \leq -\Phi/n$. To see why, note that $v(Y,Y\frac{n-1}{n}) < v(Y,Y\frac{1}{n})$ for all $Y < \overline{Y}$, which follows because the flow return function used to construct $v(Y,y)$ is strictly decreasing in $y$, and because the Markov process for $y$ is monotone. Then since $v\left(Y\,,\,\frac{Y}{n}\right) = V(Y)/n$ for all $0 \leq Y \leq \overline{Y}$, and $V(\overline{Y}) = -\Phi/n$, we obtained the desired result for $n \geq 3$. Thus if an individual country can decide to leave paying $1/n$ of the fixed cost for the entire union, the union-wide policy is not proof against individual deviations. While this theoretical
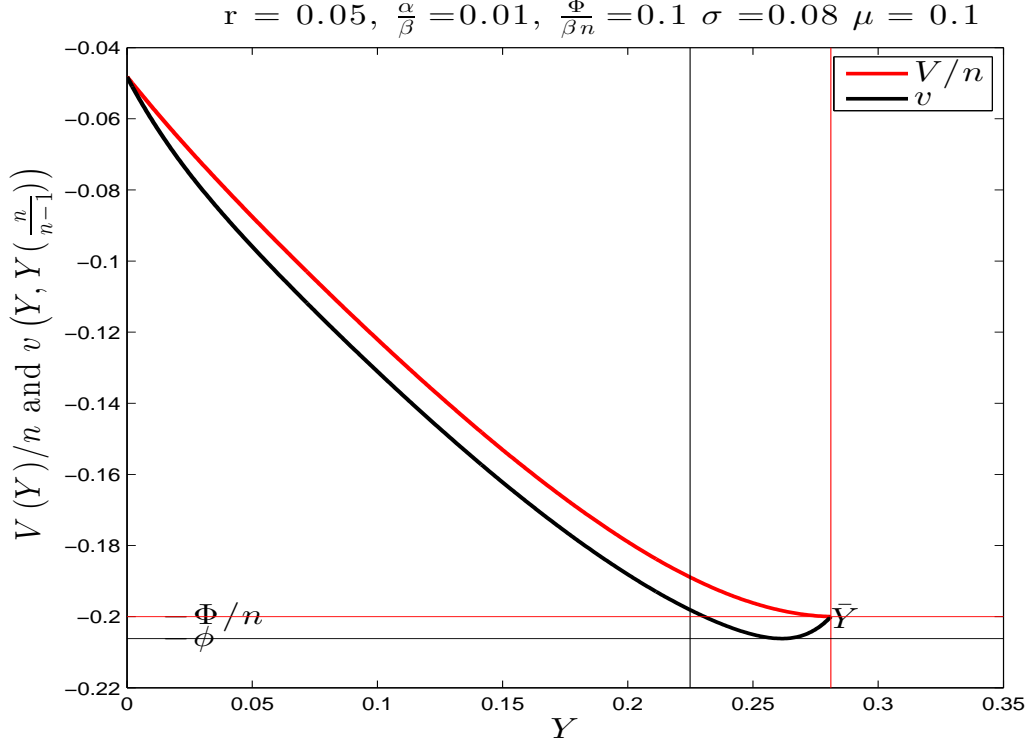
Figure 7: One country exit decision: $V/n$ vs $v$ and $\Phi/n$ vs. $\phi$

argument establishes only a weak inequality, we conjecture, and have verified for all the numerical examples that the inequality is strict for $n \geq 3$. This result is intuitive, since when $y = \overline{Y}(n-1)/n$ the misalignment of the whole union is largest in one country, and the remaining $n-1$ countries have the same misalignment (recall that the sum of the level of the misalignments is zero). If $n = 2$, then both countries' deviation will be the same, equal to the average in size but of opposite signs, and thus they will stay in the union. If $n \geq 3$ the misalignment is largest for the deviant country, and since $\overline{Y}$ makes the collective indifferent, it must make the deviant country prefer to abandon the union. We can use $v(\cdot)$ to to define $\phi$ as the smallest sum of the the fixed cost and present value of flow benefit of staying in the union for which an individual deviation so that the optimal union policy can be sustained. This can be obtained as

$$\phi = -\min\left\{v\left(y, Y\right) : 0 \leq y \leq n/(n-1)Y \,,\, 0 \leq Y \leq \overline{Y}\right\},$$

which, as we explained above, satisfies $\phi \geq \Phi/n$. The interpretation of $\phi$ in the case in which it is larger than $\Phi/n$ is that a deviant country that exits by itself will have to pay a cost higher than the pro-rated fixed cost for the collective; this difference may be a penalty

that the collective can apply to the deviant country, or the effect of an extra negative signal to investors.[8] To summarize, if there are no transfers, $\phi = \Phi/n$ for $n = 2$ an individual country will not deviate from the collective policy, while if $n \geq 3$ it requires $\phi \geq \Phi/n$ for an individual country not to deviate from the collective policy. Figure 7 illustrates this property by plotting $v(Y, Y(n-1)/n)$ and $V(Y)/n$ for the benchmark parameter values.

In Appendix M we adapt the numerical approximation of Kushner and Dupuis (2001) to solve the $v$ and compute $\phi$. In Table 3 we report $\phi$ for different configurations of parameters, around the benchmark parameter values described above. In particular, for each configuration of parameter values we $\phi$ relative to the the pro-rata of the fixed cost for the union $\Phi/n$ .

Table 3: Minimum fixed cost to deter individual country's exit: $\phi$

| $\frac{\Phi}{n}$ | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 |
|---|---|---|---|---|---|
| $\phi$ | 0.103 | 0.155 | 0.206 | 0.258 | 0.311 |
| $\beta$ | 1.0 | 1.5 | 2 | 2.5 | 3.0 |
| $\phi$ | 0.230 | 0.212 | 0.206 | 0.204 | 0.203 |
| $\sigma$ | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 |
| $\phi$ | 0.223 | 0.211 | 0.206 | 0.204 | 0.203 |

Benchmark parameter values: $r = 0.05$, $n = 5$, $\alpha = 0.02$, $\beta = 2$, $\mu = 0.1$, $\sigma = 0.08$, and $\Phi = 1.0$, so $\frac{\Phi}{n} = 0.2$. Each panel of two rows display the comparative static of $\phi$ with respect to the parameter $\Phi$, $\beta$, and $\sigma$. The middle column correspond to the benchmark parameter values.

From Table 3 we conclude that the extra fixed cost necessary to deter one country's exit is relatively small. For the benchmark case this cost is 0.206 of GDP versus a pro-rata fixed cost of $\Phi/n = 0.20$ of each country GDP, so the difference is about half a percent of yearly GDP. Looking across the wide range of parameters from Table 3, the largest value of the extra cost is about 3 percent of yearly GDP, which corresponds to the case of $\beta$, the sensitivity of flow utility to the size of the misalignment, equal to half of the benchmark value (i.e. $\beta = 1$).

# 7 Interpretation and issues left out in the analysis

In this section we review some of the important issue left out of the analysis which serves to better interpret our contribution and provide material for future research.

---

[8]Developing the setup to make the interpretations precise is beyond the scope of the present paper.

Countries in the Eurozone are highly asymmetric. We assumed symmetry of the underlying structure (although not of course of the actual realizations of random shocks) purely for reasons of tractability. In Section 2 we show the equivalence of the case of *two* asymmetric countries with symmetric countries with a geometric mean of their sensitivities $\beta$'s. This results was obtained because the union-wide monetary policy became a one dimensional case. It remains important to attempt the extension to the asymmetric situation, where the state variable will have to be a vector $\mathbf{x}$ with the components $x_i$, and the Bellman equation will be a partial differential equation. Nevertheless, we think that the message of the asymmetric case of $n = 2$ countries will be carried to some extent to the general case. We offered some conjectures about this in Section 2.

We considered a collective decision to abandon the euro completely, and one country's exit decision paying its own cost. There remains another possibility, namely that member countries that are benefitting from the euro may "force" the exit of a badly misaligned country, paying its exit cost. This requires modeling and then characterizing, as function of the state, the benefit of the union with one fewer member. Most of the time we worked with a fixed membership of the union, so a possible dependence of the parameters $\alpha_i$ and the abandonment cost $\Phi$ on the size and composition of the membership was irrelevant. Our only departure of this, the model of voluntary exit in Section 6, assumed constant returns to scale. This assumption is probably not valid, but the nature of these exit costs seems poorly understood, so it is difficult to find a clearly more realistic alternative. Furthermore, modeling this issue will require to abandon the simplification of the union problem with commitment and transfers and to model explicitly the game that is being played.

We assumed that each country $i$'s flow utility $u_i$ depends only on its own exchange rate misalignment $x_i$. In reality there can be externalities; misalignment in one country can affect other countries' flow utilities, probably negatively. However, a good model requires detailed understanding of the nature of such externalities, and that requires prior study in a different model of international macroeconomics. Indeed, the temptation of competitive devaluations, at least in a reduced form, is at the heart of Fuchs and Lippi (2006). An important issue which this will bring, and which we ignored but was was studied numerically in Fuchs and Lippi (2006), is the effect that changes in the outside option have on the policy of the union. It also makes relevant the possibility that other countries find it in their interest to "expel" a severely misaligned country, even if this requires them to bear its exit cost.[9] Using such a

---

[9]Alternatively it has been suggested that current events in the eurozone can be interpreted as small countries which find themselves bearing large costs can actually obtain transfers from the rest of the union under the threat of exit because the remaining ones fear that the exit of one will trigger a union-wide catastrophic event.

model we can compare the threshold at which such expulsion would occur to that at which the misaligned country would choose to exit voluntarily.

While our reduced form model is, as such, silent on the interpretation of the misalignment $x'$s, we have chosen to feature the traditional (old fashion?) interpretation that the cost of monetary union is the lack of country-specific monetary policy in the face of nominal rigidities and country-specific nominal shocks. Many commentators and policy makers assert that in the context of the current euro crisis sovereign defaults, bank runs, and their "contagion" also play an important role. In our model all the costs are represented by $\Phi$, and in this sense the abandonment does not feed back into the events prior to it. But alternative scenarios are conceivable and even likely. In particular our formulation of the private sector behavior does not give any important role to expectations, i.e. the law of motion of $x_i$ is taken to be independent of when abandonment occurs. An interesting and nontrivial extension will be to incorporate expectations in a way in which market dynamics generates a bank or currency run, and lead to forced exit, similar to the models of speculative attacks and currency crises such as in Krugman (1979).

More generally, we have assumed optimality of policy both before and after a break-up, and that the latter requires elimination of all the deviations from PPP. That is, we have assumed that in the Eurozone the European Central Bank acts perfectly to cancel the common component of the misalignment (see Section 2), and that after a break-up an individual country's central bank will act perfectly to offset all the deviation from PPP (so $x_i \equiv 0$). In reality monetary policy is not so perfectly conducted, and its goal may not be to eliminate such a large fraction of the PPP deviations. Indeed Mussa (1986) finds that real exchange rates tend to be more volatile when nominal exchange rates are floating. Assuming that such a large share of movements from PPP are undesirable, i.e. that the $x's$ have such a large variance, means that we are overestimating the option value. But recall, that even in that interpretation, we find relatively small magnitudes for the option value. Finally, properly modeling imperfections of policy in a convincing way is difficult and beyond what we can accomplish here.[10]

---

[10]On the theoretical side, there is a vast literature on open economy models with sticky wages and/or prices that features tractable models where coordination of monetary policy can be analyzed. See, for example, Obstfeld and Rogoff (2000), Corsetti and Pesenti (2001) and Benigno and Benigno (2003). This literature typically finds either no gains from coordinating monetary policy or small losses.

# 8 Concluding comments

Even though our model is a starkly simplified view of a welfare maximizing union, and even though it uses a reduced form representation of the private sector behavior, we argue that it has produced some new and interesting insights. It shows how a collective decision to abandon the euro can be optimal when a few countries suffer large misalignments or many countries suffer smaller misalignments. We find that for parameter values in the range consistent with conventional macro-economic views there is a non negligible but nevertheless small option value. It shows how one country's exit decision can differ from the collective decision. And at a more general level, it deepens our intuition about the effect of uncertainty on action thresholds, by emphasizing the separate effects on the threshold that would apply to a now-or-never decision and the pure option value effect. While this is a useful start, further work guided by more detailed features of the reality of the Eurozone context remains important. An actual collapse of the euro may occur and make this application redundant any day, but the general issues and methods will retain their usefulness.

# References

Alvarez, Fernando and Francesco. Lippi. 2012. "Price setting with menu cost for multi-product firms." Tech. Rep. 17923, NBER WP.

Andrade, Gregor and Steven N. Kaplan. 1998. "How Costly Is Financial (Not Economic) Distress? Evidence from Highly Leveraged Transactions That Became Distressed." *The Journal of Finance* 53 (5):pp. 1443–1493.

Benigno, Gianluca and Pierpaolo Benigno. 2003. "Price Stability in Open Economies." *The Review of Economic Studies* 70 (4):pp. 743–764.

Cecchetti, S.G., N.C. Mark, and R.J. Sonora. 2002. "Price Index Convergence Among United States Cities*." *International Economic Review* 43 (4):1081–1099.

Chevalier-Roignant, Benoît and Lenos Trigeorgis. 2012. *Competitive Strategy: Options and Games.* Cambridge, MA: MIT Prr.

Corsetti, Giancarlo and Paolo Pesenti. 2001. "Welfare and Macroeconomic Interdependence." *The Quarterly Journal of Economics* 116 (2):pp. 421–445.

Crucini, M.J. and C. Telmer. 2012. "Microeconomic sources of real exchange rate variability." Tech. rep., Tepper School of Business, Carnegie Mellon University.

Davydenko, Sergei A., Strebulaev Ilya A., and Zhao Xiaofei. 2012. "A Market-Based Study of the Cost of Default." *Review of Financial Studies* forthcoming.

Dixit, A.K. and R.S. Pindyck. 1994. *Investment under uncertainty.* Princeton Univ. Pr.

Dixit, Avinash. 1991. "Analytical Approximations in Models of Hysteresis." *Review of Economic Studies* 58 (1):141–51.

EEC. 1990. "One Market, One Money: An Evaluation of the Potential Benefits and Cost of Forming an Economic and Monetary Union." Tech. Rep. 44, European Economy.

Engel, C. and J.H. Rogers. 1996. "How Wide Is the Border?" *The American Economic Review* 86 (5):1112–1125.

———. 2001. "Deviations from purchasing power parity: causes and welfare costs." *Journal of International Economics* 55 (1):29–57.

Fuchs, W. and F. Lippi. 2006. "Monetary Union with Voluntary Participation1." *Review of Economic Studies* 73 (2):437–457.

Gali, J. and T. Monacelli. 2008. "Optimal monetary and fiscal policy in a currency union." *Journal of International Economics* 76 (1):116–132.

Karlin, S. and H.M. Taylor. 1981. *A second course in stochastic processes*, vol. 2. Academic Pr.

Krugman, P. 1979. "A model of balance-of-payments crises." *Journal of money, credit and banking* 11 (3):311–325.

Kushner, Harold J. and Paul Dupuis. 2001. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, 2nd edition.

Laeven, L. and F. Valencia. 2008. "Systemic banking crises: a new database." Tech. Rep. 2008-2224, International Monetary Fund.

Lane, P.R. 2006. "The real effects of EMU." Tech. rep., IIIS Discussion Paper No. 115.

Lothian, James R. and Mark P. Taylor. 1996. "Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries." *Journal of Political Economy* 104 (3):pp. 488–509.

Mendizabal, Hugo Rodriguez. 2002. "Monetary Union and the Transaction Cost Savings of a Single Currency." *Review of International Economics* 10 (2):263–277.

Murray, C.J. and D.H. Papell. 2002. "The purchasing power parity persistence paradigm." *Journal of International Economics* 56 (1):1–19.

Mussa, M. 1986. "Nominal exchange rate regimes and the behavior of real exchange rates: Evidence and implications." *Carnegie-Rochester Conference Series on Public Policy* 25 (1):117–214.

Obstfeld, M. and K. Rogoff. 1995. "Exchange Rate Dynamics Redux." ', *Journal of Political Economy* 103:624–60.

———. 2000. "New directions for stochastic open economy models." *Journal of international economics* 50 (1):117–153.

Rodriguez-Clare, A., C. Arkolakis, and A. Costinot. 2012. "New trade models, same old gains." *American Economic Review, forthcoming* .

Rogers, John H. 2002. "Monetary union, price level convergence, and inflation: how close is Europe to the United States?" International Finance Discussion Papers 740, Board of Governors of the Federal Reserve System (U.S.).

Rogoff, K. 1996. "The purchasing power parity puzzle." *Journal of Economic literature* 34 (2):647–668.

Rose., Andrew K. 2008. "Currency Unions." *The New Palgrave Dictionary of Economics* .

Rossi, Barbara. 2005. "Confidence Intervals for Half-Life Deviations from Purchasing Power Parity." *Journal of Business & Economic Statistics* 23 (4):pp. 432–442.

Sandleris, Guido and Mark L.J. Wright. 2011. "The Costs of Financial Crises: Resource Misallocation, Productivity and Welfare in the 2001 Argentine Crisis." Working Paper 17552, National Bureau of Economic Research.

Stokey, Nancy L. 2008. *Economics of Inaction: Stochastic Control Models with Fixed Costs.* Princeton University Press.

Strobel, F. 2007. "Joining European Monetary Union: A Real Options Perspective." *Scottish Journal of Political Economy* 54 (1):105–115.

Taylor, Alan M. and Mark P. Taylori. 2004. "The Purchasing Power Parity Debate." *The Journal of Economic Perspectives* 18 (4):135–158.

Trigeorgis, L. 1996. *Real options: Managerial flexibility and strategy in resource allocation.* the MIT Pr.

Weiss, L.A. 1990. "Bankruptcy resolution:: Direct costs and violation of priority of claims." *Journal of Financial Economics* 27 (2):285–314.

Werning, Ivan and Emmanuel Farhi. 2012. "Fiscal Unions." NBER Working Papers 18280, National Bureau of Economic Research, Inc. URL http://ideas.repec.org/p/nbr/nberwo/18280.html.

# Appendix

## A Dynamics of the state variable $Y$

$$
\begin{aligned}
dZ &= \frac{1}{n} \sum_{i=1}^{n} dX_i = \frac{1}{n} \sum_{i=1}^{n} \left[ -\mu\, X_i\, dt + \sigma\, dw_i + \sigma_c\, dw_c \right] \\
&= -\mu\, Z\, dt + \frac{\sigma}{n} \sum_{i=1}^{n} dw_i + \sigma_c\, dw_c
\end{aligned}
$$

Then

$$
\begin{aligned}
dx_i &= dX_i - dZ = -\mu\,(X_i - Z)\, dt + \sigma\, dw_i - \frac{\sigma}{n} \sum_{j=1}^{n} dw_j \qquad &\text{(A-1)} \\
&= -\mu\, x_i\, dt + \sigma\, \frac{n-1}{n}\, dw_i - \sigma\, \frac{1}{n} \sum_{j \neq i} dw_j \qquad &\text{(A-2)}
\end{aligned}
$$

Applying Itô's Lemma,

$$
\begin{aligned}
dY &= 2 \sum_{i=1}^{n} x_i\, dx_i + \tfrac{1}{2} \sum_{i=1}^{n} 2\, E[(dx_i)^2] \\
&= 2 \sum_{i=1}^{n} x_i \left[ -\mu\, x_i\, dt + \sigma\, dw_i - \frac{1}{n} \sum_{j=1}^{n} dw_j \right] + \tfrac{1}{2} \sum_{i=1}^{n} 2\, \mathrm{E}[(dx_i)^2] \\
&= -2\,\mu \sum_{i=1}^{n} x_i^2\, dt + 2 \sum_{i=1}^{n} x_i \left[ \sigma\, dw_i - \frac{\sigma}{n} \sum_{j=1}^{n} dw_j \right] \\
&\quad + n \left[ \sigma^2 \left( \frac{n-1}{n} \right)^2 dt + (n-1)\, \sigma^2\, \frac{1}{n^2} \right] dt \quad \text{using (A-2)} \\
&= \left[ (n-1)\, \sigma^2 - 2\,\mu\, Y \right]\, dt + 2\,\sigma \sum_{i=1}^{n} x_i\, dw_i - \frac{2\,\sigma}{n} \left( \sum_{i=1}^{n} x_i \right) \left( \sum_{j=1}^{n} dw_j \right) \\
&= \left[ (n-1)\, \sigma^2 - 2\,\mu\, Y \right]\, dt + 2\,\sigma \sum_{i=1}^{n} x_i\, dw_i
\end{aligned}
$$

because

$$
\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} X_i - n\, z = 0.
$$

Then

$$
E[dY^2] = 4\, \sigma^2 \sum_{i=1}^{n} x_i^2\, dt = 4\, \sigma^2\, Y\, dt \, ;
$$

therefore we can write

$$dY = \left[ (n-1) \sigma^2 - 2\mu Y \right] \, dt + 2\sigma Y^{1/2} \, dW \,,$$

where $W$ is a standard Wiener process.

# B   The case without the union-wide policy

In this section we consider the case where there is passive union-wide policy, so that $Z = 0$. Additionally we assume that there are no common shocks, i.e. $\sigma_c = 0$. In this case we have:

$$dx_i \;=\; -\mu\, x_i \, dt + \sigma \;\; dw_i \tag{A-3}$$

Applying Itô's Lemma,

$$
\begin{aligned}
dY &= 2 \sum_{i=1}^{n} x_i \, dx_i + \tfrac{1}{2} \sum_{i=1}^{n} 2\, E[(dx_i)^2] \\
&= 2 \sum_{i=1}^{n} x_i \left[ -\mu\, x_i \, dt + \sigma \;\; dw_i \right] + \tfrac{1}{2} \sum_{i=1}^{n} 2 \, \mathrm{E}[(dx_i)^2] \\
&= -2\mu \sum_{i=1}^{n} x_i^2 \, dt + 2 \sum_{i=1}^{n} x_i \sigma \, dw_i + n\,\sigma^2 dt \\
&= \left[ n\,\sigma^2 - 2\mu Y \right] \, dt + 2\sigma \sum_{i=1}^{n} x_i \, dw_i
\end{aligned}
$$

Then

$$E[dY^2] = 4\sigma^2 \sum_{i=1}^{n} x_i^2 \, dt = 4\sigma^2 Y \, dt \,,$$

therefore we can write

$$dY = \left[ \sigma^2 - 2\mu Y \right] \, dt + 2\sigma Y^{1/2} \, dW \,,$$

where $W$ is a standard Wiener process. Notice that this law of motion is identical to the case with a union-wide policy and no common shocks, except that in that case the drift features the term $n - 1$.

Indeed one can solve for the optimal stopping time by simply considering the case without a union-wide policy and changing two parameters, which we now label with $'$, namely the number of countries and the constant in the instantaneous return function. In particular, the optimal value of $\overline{Y}$ is the same in the original problem with the union wide policy $Z$ in equation (6) and in the problem with: *i)* $Z = 0$, *ii)* no aggregate shocks, i.e. $\sigma_c = 0$, *iii)* the number of countries equal to $n' = n - 1$, and *iv)* the value $\alpha' = \alpha\, n/(n-1)$. Due to  *i)* and

*ii)* shocks to each country misalignment are independent, but with $Z = 0$, the process for $Y$ is more volatile, i.e. its drift is larger in algebraic value, as shown above. The adjustment in *iv)* is because the problem still has the instantaneous return function $U = n\alpha - \beta/2\,Y$.

# C   The case of two asymmetric countries

We can write

$$Z = \frac{\beta_1}{\beta_1 + \beta_2}X_1 + \frac{\beta_2}{\beta_1 + \beta_2}X_2 = \arg\min_z \beta_1(X_1 - z)^2 + \beta_2(X_2 - z)^2$$

The minimized value satisfy

$$\beta_1(X_1 - Z)^2 + \beta_2(X_2 - Z)^2 = \frac{\beta_1\beta_2^2 + \beta_2\beta_1^2}{(\beta_1 + \beta_2)^2}(X_1 - X_2)^2 = \frac{\beta_1\beta_2}{\beta_1 + \beta_2}(X_1 - X_2)^2$$

Note that if $\beta_1 = \beta_2 = \beta$ then

$$\beta\,Y = \beta_1(X_1 - Z)^2 + \beta_2(X_2 - Z)^2 = \frac{\beta}{2}(X_1 - X_2)^2$$

which suggest to define:

$$Y \equiv (X_1 - X_2)^2/2 \,. \tag{A-4}$$

To derive an expression for $dY$ we first note that

$$d(X_1 - X_2) = -\mu\,(X_1 - X_2)\,dt + \sqrt{\sigma_1^2 + \sigma_2^2}\,dW,$$

and using Ito's lemma and the definition of $Y$ we obtain:

$$dY = \left[\left(\frac{\sigma_1^2 + \sigma_2^2}{2}\right) - \mu\,2Y\right]dt + 2\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}\,Y}\,dW \tag{A-5}$$

Note that for the case of $\sigma_1 = \sigma_2 = \sigma$ this expression is the same as the benchmark model with homogeneous countries and with $n = 2$.

Thus the expression for $\overline{Y}$ for the economy with heterogenous $\alpha$'s, $\beta$'s and $\sigma$'s is the same as it will be obtained if the union has homogenous countries with the triplet $(\bar{\alpha}, \bar{\sigma}, \bar{\beta})$ given by:

$$\bar{\alpha} = \frac{\alpha_1 + \alpha_2}{2} \,, \ \bar{\sigma}^2 = \frac{\sigma_1^2 + \sigma_2^2}{2} \ \text{ and } \ \bar{\beta} = 2\frac{\beta_1\,\beta_2}{\beta_1 + \beta_2} \le \frac{\beta_1 + \beta_2}{2} \,, \tag{A-6}$$

with equality if and only if $\beta_1 = \beta_2$. This inequality follows because $\bar{\beta}$ is the harmonic mean of $\beta_1$ and $\beta_2$. With this definitions we can write the flow utility as:

$$u_1 + u_2 \equiv \alpha_1 + \alpha_2 + \beta_1(X_1 - Z)^2 + \beta_2(X_2 - Z)^2 = n\,\bar{\alpha} + \bar{\beta}\,Y \tag{A-7}$$

34

where $Y$ is defined by equation (A-4) and follows:

$$dY = \left[(n-1)\bar{\sigma}^2 - 2\,\mu\,Y\right] dt + 2\,\sqrt{\bar{\sigma}}\,Y\,dW \;, \tag{A-8}$$

for $n = 2$.

Since we show that $\overline{Y}$ is increasing in $\bar{\beta}$ in the homogenous case, dispersion on the $\beta$'s decrease $\overline{Y}$, because the harmonic mean is smaller than the arithmetic mean.

## D  Power series solution

Here we derive the power series solution to the differential equation (9). Substituting (11) into it, we have

$$
\begin{aligned}
0 \;=\;& 2\,\sigma^2\,Y\sum_{m=2}^{\infty} m(m-1)\,c_m\,Y^{m-2} + \left[(n-1)\,\sigma^2 - 2\,\mu\,Y\right]\sum_{m=1}^{\infty} m\,c_m\,Y^{m-1} \\
& -\,r\sum_{m=0}^{\infty} c_m\,Y^m + n\,\alpha - \tfrac{1}{2}\,\beta\,Y \\
\;=\;& 2\,\sigma^2\sum_{m=2}^{\infty} m(m-1)\,c_m\,Y^{m-1} + (n-1)\,\sigma^2\sum_{m=1}^{\infty} m\,c_m\,Y^{m-1} - 2\,\mu\sum_{m=1}^{\infty} m\,c_m\,Y^m \\
& -\,r\sum_{m=0}^{\infty} c_m\,Y^m + n\,\alpha - \tfrac{1}{2}\,\beta\,Y \\
\;=\;& 2\,\sigma^2\sum_{m'=1}^{\infty}(m'+1)m'\,c_{m'+1}\,Y^{m'} + (n-1)\,\sigma^2\sum_{m'=0}^{\infty}(m'+1)\,c_{m'+1}\,Y^{m'} - 2\,\mu\sum_{m=1}^{\infty} m\,c_m\,Y^m \\
& -\,r\sum_{m=0}^{\infty} c_m\,Y^m + n\,\alpha - \tfrac{1}{2}\,\beta\,Y \\
\;=\;& \left[(n-1)\,\sigma^2\,c_1 - r\,c_0 + n\,\alpha\right] + \left[4\,\sigma^2\,c_2 + 2\,(n-1)\,\sigma^2\,c_2 - 2\,\mu\,c_1 - r\,c_1 - \tfrac{1}{2}\,\beta\right]Y \\
& +\sum_{m=2}^{\infty}\left[2\,\sigma^2\,m(m+1)\,c_{m+1} + (n-1)\,\sigma^2\,(m+1)\,c_{m+1} - 2\,\mu\,m\,c_m - r\,c_m\right]Y^m \\
\;=\;& \left[(n-1)\,\sigma^2\,c_1 - r\,c_0 + n\,\alpha\right] + \left[2\,(n+1)\,\sigma^2\,c_2 - (2\,\mu+r)\,c_1 - \tfrac{1}{2}\,\beta\right]Y \\
& +\sum_{m=2}^{\infty}\left[(2\,m+n-1)\,(m+1)\,\sigma^2\,c_{m+1} - (2\,\mu\,m+r)\,c_m\right]Y^m
\end{aligned}
$$

As this is an identity in $Y$, we equate the coefficients of all powers of $y$ separately to zero, yielding

$$c_1 = \frac{r\,c_0 - n\,\alpha}{(n-1)\,\sigma^2}\,, \qquad c_2 = \frac{\tfrac{1}{2}\,\beta + (2\,\mu+r)\,c_1}{2\,(n+1)\,\sigma^2}$$

and

$$c_{m+1} = \frac{2\,\mu\,m+r}{(2\,m+n-1)\,(m+1)\,\sigma^2}\,c_m \quad \text{for all } m \geq 2.$$

Since $c_{m+1}/c_m \to 0$ as $m \to \infty$, the series converges absolutely for all $Y$.

# E    Now-or-never decision

For each country, we have (A-2),

$$dx_i(t) = -\mu \, x_i(t) \, dt + \sigma \, \frac{n-1}{n} \, dw_i - \sigma \, \frac{1}{n} \sum_{j \neq i} dw_j \, .$$

Therefore

$$d\left[x_i(t) \, e^{\mu t}\right] = e^{\mu t} \left[dx_i(t) + \mu \, x_i(t)\right] = \sigma \, e^{\mu t} \left[\sigma \, \frac{n-1}{n} \, dw_i - \sigma \, \frac{1}{n} \sum_{j \neq i} dw_j\right] ,$$

and

$$x_i(t) \, e^{\mu t} = x_i(0) + \sigma \int_0^t e^{\mu s} \left[\sigma \, \frac{n-1}{n} \, dw_i - \sigma \, \frac{1}{n} \sum_{j \neq i} dw_j\right]$$

or

$$x_i(t) = x_i(0) \, e^{-\mu t} + \sigma \int_0^t e^{-\mu (t-s)} \left[\sigma \, \frac{n-1}{n} \, dw_i - \sigma \, \frac{1}{n} \sum_{j \neq i} dw_j\right]$$

Therefore

$$
\begin{aligned}
\mathrm{E}_0\left[x_i(t)^2\right] &= x_i(0)^2 \, e^{-2\mu t} + \sigma^2 \left[\left(\frac{n-1}{n}\right)^2 + (n-1)\left(\frac{1}{n}\right)^2\right] \int_0^t e^{-2\mu (t-s)} \, ds \\[2mm]
&= x_i(0)^2 \, e^{-2\mu t} + \sigma^2 \, \frac{n-1}{n} \, \frac{1 - e^{-2\mu t}}{2\mu} \, .
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\int_0^T \mathrm{E}_0\left[x_i(0)^2\right] e^{-rt} \, dt &= x_i(0)^2 \, \frac{1 - e^{-(2\mu+r)T}}{2\mu+r} + \frac{\sigma^2}{2\mu} \left\{\frac{1 - e^{-rT}}{r} - \frac{1 - e^{-(2\mu+r)T}}{2\mu+r}\right\} \\[2mm]
&= \left[x_i(0)^2 - \frac{n-1}{n} \, \frac{\sigma^2}{2\mu}\right] \frac{1 - e^{-(2\mu+r)T}}{2\mu+r} + \frac{n-1}{n} \, \frac{\sigma^2}{2\mu} \, \frac{1 - e^{-rT}}{r} \, .
\end{aligned}
$$

Summing over $i$,

$$\int_0^T \mathrm{E}_0\left[Y(t)\right] e^{-rt} \, dt = \left[Y(0) - \frac{(n-1)\sigma^2}{2\mu}\right] \frac{1 - e^{-(2\mu+r)T}}{2\mu+r} + \frac{(n-1)\sigma^2}{2\mu} \, \frac{1 - e^{-rT}}{r} \, .$$

Then the expected aggregate utility from continuation of the euro until time $T$ is

$$
\begin{aligned}
&\int_0^T \left(n\alpha - \tfrac{1}{2}\beta \, \mathrm{E}_0\left[Y(t)\right]\right) e^{-rt} \, dt \\[2mm]
&= n\alpha \, \frac{1 - e^{-rT}}{r} - \tfrac{1}{2}\beta \left\{\left[Y(0) - \frac{(n-1)\sigma^2}{2\mu}\right] \frac{1 - e^{-(2\mu+r)T}}{2\mu+r} + \frac{(n-1)\sigma^2}{2\mu} \, \frac{1 - e^{-rT}}{r}\right\} \\[2mm]
&= \left[n\alpha - \frac{(n-1)\beta\sigma^2}{4\mu}\right] \frac{1 - e^{-rT}}{r} - \tfrac{1}{2}\beta \left[Y(0) - \frac{(n-1)\sigma^2}{2\mu}\right] \frac{1 - e^{-(2\mu+r)T}}{2\mu+r}
\end{aligned}
$$

Letting $T \to \infty$, we have the value of starting at $Y(0)$ and continuing for ever:

$$
\begin{aligned}
V_E(Y(0)) &= \int_0^\infty \left( n\,\alpha - \tfrac{1}{2}\,\beta\,E_0\left[Y(t)\right] \right) e^{-rt}\, dt \\
&= \left[ n\,\alpha - \frac{(n-1)\,\beta\,\sigma^2}{4\,\mu} \right] \frac{1}{r} - \tfrac{1}{2}\,\beta \left[ Y(0) - \frac{(n-1)\,\sigma^2}{2\,\mu} \right] \frac{1}{2\,\mu + r} \quad \text{(A-9)}
\end{aligned}
$$

If the only choice is to abandon either now or never, it is better to abandon if $V_E(Y(0)) < -\Phi$, i.e.

$$
\tfrac{1}{2}\,\beta \left[ Y(0) - \frac{(n-1)\,\sigma^2}{2\,\mu} \right] \frac{1}{2\,\mu + r} > \Phi + \left[ n\,\alpha - \frac{(n-1)\beta\,\sigma^2}{4\,\mu} \right] \frac{1}{r},
$$

or

$$
\begin{aligned}
Y(0) \;>\; & \frac{(n-1)\,\sigma^2}{2\,\mu} + \frac{2\,(2\mu + r)}{\beta}\,\Phi + \frac{2\,(2\mu + r)}{\beta} \left[ n\,\alpha - \frac{(n-1)\,\beta\,\sigma^2}{4\,\mu} \right] \frac{1}{r} \\
&= \frac{2\,(2\mu + r)}{r\,\beta}\,[\,r\,\Phi + n\,\alpha\,] + (n-1)\left[ 1 - \frac{2\,\mu + r}{r} \right] \frac{\sigma^2}{2\,\mu} \\
&= \frac{2\,(2\mu + r)}{r\,\beta}\,[\,r\,\Phi + n\,\alpha\,] - \frac{(n-1)\,\sigma^2}{r}
\end{aligned}
$$

The right hand side in the last line is the threshold $\widehat{Y}$ for the now-or-never abandonment decision.

# F   Comparative Static of $\bar{Y}$

In this appendix we show four comparative static results of the optimal threshold $\bar{Y}$ sated in Section 4.2. We start with a result we use to prove this claim. Let $\theta$ be any of the parameters $\alpha, \mu$, or $\beta$. Let $\bar{Y}(\theta)$ be the optimal threshold as a function of the parameter. Fix a value $Y' \geq 0$ and let $V(Y; Y', \theta)$ be defined for all $Y \leq Y'$:

$$
V\left(Y\,;\,Y',\,\theta\right) = \mathrm{E}\left[ \left. \int_0^{\tau(Y')} U\left(Y(t)\,;\,\theta\right)\, e^{-rt}\, dt - e^{-r\,\tau(Y')}\,\Phi \;\right|\; Y(0) = Y\,,\,\theta \right] \quad \text{(A-10)}
$$

and $V\left(Y\,;\,Y',\,\theta\right) = -\Phi$ for $Y > Y'$, where $\tau(Y')$ is the first time that $Y(t)$ hits $Y'$.

LEMMA 1.   Assume that

$$
V\left(Y\,;\,\bar{Y}(\theta),\,\theta'\right) > V\left(Y\,;\,\bar{Y}(\theta),\,\theta\right) \quad \text{for all } Y \leq \bar{Y}(\theta),
$$

with equality if $Y = \bar{Y}(\theta)$. Then $\bar{Y}(\theta') \geq \bar{Y}(\theta)$.

To prove this lemma assume that $\bar{Y}(\theta') < \bar{Y}(\theta)$. Consider to follow the policy $\tau(\bar{Y}(\theta))$ when the value of the parameter is $\theta'$. By hypothesis

$$
-\Phi = V\left(\bar{Y}(\theta')\,;\,\bar{Y}(\theta'),\,\theta'\right) < V\left(\bar{Y}(\theta')\,;\,\bar{Y}(\theta),\,\theta'\right)
$$

a contradiction with the optimality of $\overline{Y}(\theta')$ and $\overline{Y}(\theta') < \overline{Y}(\theta)$.

To show 1 we note that $U(Y, \alpha)$ is strictly increasing in $\alpha$ for each $Y$, and thus we verify the hypothesis of the lemma. To show 4 we note that the distribution of $Y(t)$ conditional on $Y(0) = Y$ and $Y(s) < Y'$ for all $s \in (0, t)$ is stochastically lower for higher values of $\mu$. Since $U(Y)$ is decreasing in $Y$ and $\Phi > 0$ then we verify the hypothesis of the lemma. To show 2 we note that $U(Y, \alpha)$ is strictly decreasing in $\beta$ for each $Y > 0$, and thus we verify the hypothesis of the lemma. To show 3 we note that $\overline{Y}$ depends only on the sum $r\Phi + n\alpha$ and hence, it follows from 1. To show 5 compute $\partial V(Y; \overline{Y})/\partial r$ evaluate it at $Y < \overline{Y}$. When $\alpha = 0$ this expression is strictly positive if $\beta > 0$.

# G  The case of small discount rates

In this section we analyze the case where $r = 0$, which allows some analytical comparative statics. This also provides a good approximation to the solution when $r > 0$ provided that $\sigma^2$ is large relative to $r$ and the range of inaction is small. Furthermore, we use this case to develop an approximation for small value of $\mu$ and $r$. The result of this section are developed for the case with $Z = 0$, and hence $n$ and $\alpha$ has to be changed accordingly as explained in Appendix B.

In the undiscounted case the interesting set of parameters is the one for which $\overline{Y} < \infty$ and the corresponding stopping time is finite. This set is characterized by $\beta/\alpha > 4\mu/\sigma^2$. Furthermore, for $r = 0$ we set $\Phi = 0$ since in the case where the thresholds $\overline{Y}$ is achieved with probability one, the fixed cost has no effect on the problem.

With these parameters, we find that the threshold $\overline{Y}$:

i) is only a function of $n$ and the ratios $\beta/\alpha$ and $\mu/\sigma^2$,

ii) is strictly decreasing in $\beta/\alpha$,

iii) tends to 0 as $\beta/\alpha \to \infty$,

iv) tends to $+\infty$ as $\beta/\alpha \to \infty$,

v) is given by $\overline{Y} = 2(n+2)\alpha/\beta$ which is independent of $\sigma^2$ when $\mu = 0$, and

vi) is (locally) increasing in $\mu/\sigma^2$ when evaluated near $\mu/\sigma^2 = 0$, with derivative: $16\,(\alpha/\beta)^2\,(2+n)/(4+n)$ at zero.

For the case of low discounting (small $r$ and $\mu$), we have an approximation for the threshold $\overline{Y}$:

$$\overline{Y} = 2(n+2)\left(\frac{\alpha}{\beta} + \frac{r\,\Phi/n}{\beta}\right) + 16\left(\frac{n+2}{n+4}\right)\left(\frac{\alpha}{\beta}\right)^2\left(\frac{\mu - r/n}{\sigma^2}\right) + o\left(||(\mu, r)||\right)$$

where $o(x)$ denote terms of order smaller than $x$. The expression in equation (14) is obtained by using the equivalence for the case where $Z = 0$, as explained in Appendix B.

The proofs are in Appendix H.

# H   Propositions for zero and low discount rates

For notational purposes this appendix uses the case where there the collective is not using the "optimal union-wide policy", i.e. when $Z = 0$ and $\sigma_c = 0$. Appendix B explains how to map the results for the case with $Z = 0$ into the case with a optimal union-wise policy as in equation (6).

For $r = 0$ the problem becomes:

$$V(Y) = E\left[\int_0^\tau U\left(Y\left(t\right)\right)\, dt \mid Y(0) = Y\right]$$

where

$$U(Y) = n\,\alpha - \frac{\beta}{2}Y\ .$$

First we characterize the value of inaction, i.e. the value that results from setting $\tau = \infty$.

**Proposition 1:** Assume $r = 0$ and $\sigma^2 > 0$. Let $\tau = \infty$ and denote the value of staying forever in the union as $\hat{V}$. If $\mu > 0$:

$$\hat{V}(Y) = \begin{cases} +\infty & \text{if } \frac{\beta}{\alpha} < 4\frac{\mu}{\sigma^2} \\ \frac{1}{4}\frac{\beta}{\mu}\left[n\frac{\sigma^2}{2\mu} - Y(0)\right] & \text{if } \frac{\beta}{\alpha} = 4\frac{\mu}{\sigma^2} \\ -\infty & \text{if } \frac{\beta}{\alpha} > 4\frac{\mu}{\sigma^2} \end{cases}$$

If $\mu = 0$, then $\hat{V}(Y) = -\infty$.

**Proof of Proposition 1:**

Each country's $x$ follows:

$$x_i(t) = x_i(0)e^{-\mu t} + \sigma \int_0^t e^{-\mu(t-s)}dW_{i,s}$$

thus

$$\mathrm{E}_0\left[x_i(t)^2\right] = x_i(0)^2 e^{-2\mu t} + \sigma^2 \int_0^t e^{-2\mu(t-s)}ds = x_i(0)^2 e^{-2\mu t} + \sigma^2 \frac{1 - e^{-2\mu t}}{2\mu}$$

and integrating this expression we obtain up to a fixed time $T > 0$:

$$\begin{aligned} \int_0^T \mathrm{E}_0\left[x_i(t)^2\right]dt &= x_i(0)^2 \frac{1 - e^{-2\mu T}}{2\mu} + \frac{\sigma^2}{2\mu}T - \sigma^2 \int_0^T \frac{e^{-2\mu t}}{2\mu}dt \\ &= x_i(0)^2 \frac{1 - e^{-2\mu T}}{2\mu} + \frac{\sigma^2}{2\mu}T - \sigma^2 \frac{1 - e^{-2\mu T}}{4\mu^2} \end{aligned}$$

Note that $Y(t) = \sum_i^n x_i(t)^2$ and consider

$$
\begin{aligned}
\hat{V}(Y;T) &\equiv \int_0^T \mathrm{E}\left[U\left(Y\left(t\right)\right) \mid Y(0) = Y\right] dt = n\alpha T - \frac{\beta}{2}\mathrm{E}\left[Y(t) | Y(0) = Y\right] \\
&= n\,\alpha\,T - \frac{\beta}{2}\sum_{i=1}^n \int_0^T \mathrm{E}_0\left[x_i(t)^2\right] dt \\
&= nT\left(\alpha - \frac{\beta\sigma^2}{4\mu}\right) - \frac{\beta}{2}\left(Y(0)\frac{1-e^{-2\mu T}}{2\mu} + n\sigma^2\frac{1-e^{-2\mu T}}{4\mu^2}\right)
\end{aligned}
$$

Taking $T \to \infty$ we obtain the desired expressions. If $\mu = 0$ we have $\mathrm{E}_0[Y(t)] = Y(0) + n\sigma^2 t$ so that:

$$
\hat{V}(Y;T) \equiv\ = n\alpha\,T - \frac{\beta}{2}\left(Y(0)\,T + n\sigma^2\frac{T^2}{2}\right)
$$

We are interested in the configuration of parameters for which $\hat{V} = -\infty$, i.e. the case where $\beta/\alpha > 4\mu/\sigma^2$. In this case the optimal decision rule will involve exiting the union when a $\bar{Y} < \infty$ is reached. We now show that the value of such a policy for any $\bar{Y} < \infty$ is finite. A preliminary step is to show that the stopping time for an arbitrary threshold is finite.

**Proposition 2:** Assume $r = 0$ and $\sigma^2 > 0$. Let $\tau(\bar{Y})$ denote the first time $Y$ reaches $\bar{Y} < \infty$. This stopping time is finite with probability one, and:

$$
\mathrm{E}\left[\tau(\bar{Y}) \mid Y(0) = Y\right] = \frac{1}{n\sigma^2}(\bar{Y} - Y) + \frac{1}{(n+2)n\sigma^2}\frac{\mu}{\sigma^2}\sum_{m=2}^\infty D_m\left[\frac{\mu}{\sigma^2}\right]^{m-2}\left(\bar{Y}^m - Y^m\right) < \infty
$$

where

$$
D_m \equiv \prod_{j=2}^{m-1} d_j \text{ for } m \geq 3, \quad D_2 = 1, \quad \text{and} \quad d_m \equiv \frac{2m}{(2m+n)(m+1)} \text{ for } m \geq 2. \quad \text{(A-11)}
$$

**Proof of Proposition 2:**

Fix $0 < \bar{Y} < \infty$. Let $\mathcal{T}(Y) = \mathrm{E}\left[\tau(\bar{Y}) \mid Y(0) = Y\right]$. This function satisfies the same ordinary differential equation as the Bellman equation in the inaction region, with $\beta = r = 0$ and $\alpha = 1/n$. Thus it has a power series representation with coefficients:

$$
\mathcal{T}(Y) = c_0 - \frac{1}{n\sigma^2}Y - \frac{1}{(n+2)n\sigma^2}\frac{\mu}{\sigma^2}\sum_{m=2}^\infty D_m\left[\frac{\mu}{\sigma^2}\right]^{m-2}Y^m
$$

40

The boundary condition is that $\mathcal{T}(\bar{Y}) = 0$ so we solve for $c_0$ to get:

$$\mathcal{T}(Y) = \frac{1}{n\sigma^2}(\bar{Y} - Y) + \frac{1}{(n+2)n\sigma^2}\frac{\mu}{\sigma^2}\sum_{m=2}^{\infty} D_m \left[\frac{\mu}{\sigma^2}\right]^{m-2} (\bar{Y}^m - Y^m).$$

Note that, as required, $\mathcal{T}(Y) > 0$ for $Y \in [0, \bar{Y})$.

In the case of $\mu = 0$ the expected time until reaching the barrier $\bar{Y}$ starting from $Y$ is simply: $(\bar{Y} - Y)/(n\sigma^2)$, a result obtained in Alvarez and Lippi (2012).

Proposition 2 implies that, for the parameter for which the value of never exiting the union diverges to $-\infty$, the value of of following a threshold policy is finite, since for any $\bar{Y} < \infty$:

$$V(Y) \geq \left(n\alpha - \frac{\beta}{2}\bar{Y}\right) E\left[\tau(\bar{Y}) \,|\, Y(0) = Y\right] > -\infty \,,$$

and moreover, $V(0) > 0$, by considering the feasible policy $\bar{Y} = n\,\alpha/\beta$ and using the previous lower bound.

The next step is to find the optimal threshold $\bar{Y}$ as a function of the parameters. We do so by using the power series solution of the differential equation for the value function in the range of inaction given in the appendix and imposing smooth pasting. This gives one equation in one unknown.

**Proposition 3:** Assume that $\sigma^2 > 0$, $r = 0$, and that $\beta/\alpha > 4\mu/\sigma^2$. Then the optimal threshold $\bar{Y}$ is finite, and it is given by the unique positive solution to:

$$1 = \frac{1}{4(n+2)}\left[\frac{\beta}{\alpha} - 4\frac{\mu}{\sigma^2}\right]\sum_{m=2}^{\infty} D_m\,m\left[\frac{\mu}{\sigma^2}\right]^{m-2} \bar{Y}^{m-1} \,,$$

where $D_m$'s are given in (A-11). Moreover the threshold $\bar{Y}$: i) is only a function of $n$ and the ratios $\beta/\alpha$ and $\mu/\sigma^2$, ii) is strictly decreasing in $\beta/\alpha$, iii) tends to 0 as $\beta/\alpha \to \infty$, iv) tends to $+\infty$ as $\beta/\alpha \to \infty$, v) is given by $\bar{Y} = 2(n+2)\alpha/\beta$ which is independent of $\sigma^2$ when $\mu = 0$, and vi) is (locally) increasing in $\mu/\sigma^2$ when evaluated near $\mu/\sigma^2 = 0$ .

**Proof of Proposition 3:**

The power series representation for the solution of the ordinary differential equation from Appendix A gives, for the case when $r = 0$, the following expression for $V'(Y)$:

$$V'(Y) = c_1 + \sum_{m=2}^{\infty} c_m\,m\,Y^{m-1} = -\frac{\alpha}{\sigma^2} + \frac{\alpha}{4(n+2)\sigma^2}\left[\frac{\beta}{\alpha} - 4\frac{\mu}{\sigma^2}\right]\sum_{m=2}^{\infty} D_m\,m\left[\frac{\mu}{\sigma^2}\right]^{m-2} Y^{m-1}$$

41

Setting $V'(\bar{Y}) = 0$ we obtain:

$$1 = \frac{1}{4(n+2)} \left[ \frac{\beta}{\alpha} - 4\frac{\mu}{\sigma^2} \right] \sum_{m=2}^{\infty} D_m \, m \left[ \frac{\mu}{\sigma^2} \right]^{m-2} \bar{Y}^{m-1}$$

where the $D_m$ coefficients were defined in Proposition 2.

Since $V'(0) = -\alpha/\sigma^2 < 0$, then the smallest strictly positive value $\bar{Y} > 0$ for which $V'(\bar{Y}) = 0$ must be a local minimum. By hypothesis $\frac{\beta}{\alpha} - 4\frac{\mu}{\sigma^2} > 0$, and thus the right hand side of this expression goes to zero as $\bar{Y}$ goes to zero and goes to infinity as $\bar{Y}$ diverges, hence there is always a unique strictly solution to this equation.

By inspection the implied solution for $\bar{Y}$ only depends on $n$, $\beta/\alpha$, and $\mu/\sigma^2$.

Since the right hand side is increasing in increasing in $\beta/\alpha$ for any $\bar{Y}$, then the optimal threshold is decreasing. The limits cases for $\bar{Y}$ follow directly from this argument.

The expression for the case of $\mu/\sigma^2 = 0$ follows from direct computation.

To obtain $\partial \bar{Y} / \partial(\mu/\sigma^2) < 0$ at $\mu/\sigma^2 = 0$ we set $\alpha = \sigma^2 = 1$ (which is without loss of generality), normalize the right hand side by $4(n+2)$ and differentiate it w.r.t. to $\mu$ and evaluate at $\mu = 0$ obtaining:

$$-4D_2 \times 2\bar{Y} + \beta D_3 \, 3 \, \bar{Y}^2 = \bar{Y} \left[ -8 + \beta d_2 \, 3 \, \bar{Y} \right] = \bar{Y} \left[ -8 + \beta \frac{4}{(4+n)3} \, 3 \, \bar{Y} \right]$$

where the last lines use the expressions for $D_3$ and for $d_2$. Replacing the value of $\bar{Y}$ for $\mu = 0$ we obtain

$$\begin{aligned} -8 \, D_2 \, \bar{Y} + \beta D_3 \, 3 \, \bar{Y}^2 &= 4\bar{Y} \left[ -2 + \frac{\beta}{(4+n)} \frac{2(n+2)}{\beta} \right] \\ &= 4\bar{Y} \left[ -2 + \frac{4+2n}{4+n} \right] = -\frac{16 \, \bar{Y}}{4+n} < 0 \end{aligned}$$

This gives a strictly negative derivative at $\mu = 0$. Since the RHS is $C^1$ w.r.t. to $\mu/\sigma^2$ and $\bar{Y}$ and since its derivative w.r.t. $\bar{Y}$ is non-zero, the implicit function theorem implies that $\bar{Y}$ is strictly increasing in $\mu/\sigma^2$ a neighborhood of $\mu/\sigma^2 = 0$. To obtain $\partial \bar{Y} / \partial(\mu/\sigma^2) < 0$ at $\mu/\sigma^2 = 0$ let the RHS of equation determining $\bar{Y}$ be $G(u, \bar{Y})$ where $u \equiv \mu/\sigma^2$. We need to compute $G_u(0, \bar{Y})$ and $G_{\bar{Y}}(0, \bar{Y})$ and evaluate it at $\bar{Y} = (\alpha/\beta)2(n+2)$. We have:

$$4(n+2) \, G_u(u, \bar{Y}) = -4 \sum_{m=2}^{\infty} D_m \, m \, [u]^{m-2} \, \bar{Y}^{m-1} + \left[ \frac{\beta}{\alpha} - 4u \right] \sum_{m=2}^{\infty} D_m \, m \, (m-2) \, [u]^{m-3} \, \bar{Y}^{m-1}$$

Evaluated at $u = 0$ we have:

$$4(n+2) \, G_u(0, \bar{Y}) = -4 \, D_2 \, 2\bar{Y} + \frac{\beta}{\alpha} D_3 \, 3 \, \bar{Y}^2$$

We also have:

$$4(n+2)\, G_{\bar{Y}}(u, \bar{Y}) = \left[\frac{\beta}{\alpha} - 4u\right] \sum_{m=2}^{\infty} D_m\, m\,(m-1)\, [u]^{m-2}\, \bar{Y}^{m-2}$$

which evaluated at $u = 0$ gives:

$$4(n+2)\, G_{\bar{Y}}(0, \bar{Y}) = \frac{\beta}{\alpha}\, D_2\, 2$$

Thus,

$$\begin{aligned}
\frac{\partial \bar{Y}(\mu/\sigma^2)}{\partial(\mu/\sigma^2)} &= -\frac{-4\, D_2\, 2\bar{Y} + \frac{\beta}{\alpha} D_3\, 3\, \bar{Y}^2}{\frac{\beta}{\alpha} D_2\, 2} = -\bar{Y}\left[-4\frac{\alpha}{\beta} + \frac{D_3}{D_2}\frac{3}{2}\bar{Y}\right] \\
&= -\bar{Y}\left[-4\frac{\alpha}{\beta} + \frac{4}{(4+n)3}\frac{3}{2}\bar{Y}\right] = -\bar{Y}4\left[-\frac{\alpha}{\beta} + \frac{1}{(4+n)2}\bar{Y}\right]
\end{aligned}$$

and replacing $\bar{Y} = (\alpha/\beta)2(n+2)$ we have:

$$\begin{aligned}
\frac{\partial \bar{Y}(\mu/\sigma^2)}{\partial(\mu/\sigma^2)} &= -2(n+2)\left(\frac{\alpha}{\beta}\right)^2 4\left[-1 + \frac{2(n+2)}{(4+n)2}\right] = 2(n+2)\left(\frac{\alpha}{\beta}\right)^2 4\frac{2}{4+n} \\
&= 16\left(\frac{\alpha}{\beta}\right)^2 \frac{2+n}{4+n}\,.
\end{aligned}$$

The next proposition provides an approximation for the threshold $\bar{Y}$ for small values of $\mu$ and $r$.

**Proposition 4:** Assume that $\beta > 0$, $\alpha > 0$, $\sigma^2 > 0$ and that $r, \mu, \Phi$ are non-negative and that $\beta/\alpha > 4\mu/\sigma^2$. The optimal threshold $\bar{Y}$ satisfies

$$\bar{Y} = 2(n+2)\left(\frac{\alpha}{\beta} + \frac{r\,\Phi/n}{\beta}\right) + 16\left(\frac{n+2}{n+4}\right)\left(\frac{\alpha}{\beta}\right)^2\left(\frac{\mu - r/n}{\sigma^2}\right) + o\left(||(\mu, r)||\right)$$

where $o(x)$ denote of order smaller than $x$.

Proposition 4 is based on a first order expansion of $\bar{Y}$ around $(r, \mu)$ around $(0, 0)$ for fixed strictly positive value of $\beta, \alpha, \sigma^2$, i.e.:

$$\bar{Y}(\mu, r) = \bar{Y}(0, 0) + \frac{\partial \bar{Y}(0, 0)}{\partial r} r + \frac{\partial \bar{Y}(0, 0)}{\partial \mu} \mu + o(||(\mu, r)||)\,.$$

The approximate expression in equation (A-12) shows that whether $\bar{Y}$ is increasing (resp. decreasing) with $\sigma^2$ or not depends on comparing whether $r/n - \mu$ is positive (resp. negative). Note that when $r = n\,\mu$ the threshold is independent of $\sigma^2$, generalizing the result obtained for $r = \mu = 0$. Note that when $r/n \neq \sigma^2$, the value of $\bar{Y}$ can be very sensitive w.r.t. $\sigma^2$ if the volatility is very small.

Regardless of whether $\bar{Y}$ is increasing or decreasing on $\sigma^2$, on the domain of Proposition 4, i.e. when $\beta/\alpha > 4\mu/\sigma^2$, it is easy to show that $\bar{Y} - \hat{Y}$, the difference between the optimal threshold and the "now or never" threshold is increasing on $\sigma^2$. In other words, the option value is increasing in $\sigma^2$.

Finally, while the approximation is only valid for small $r$ and $\mu$, it accurately predicts whether $\bar{Y}$ is monotone or not. Note that for $n = 5$, $r = 0.05$ and $\mu = 0.01$ it accurately predicts that $\bar{Y}$ does not depend on $\sigma^2$. Moreover, for the values $\Phi = 100$, $\beta = 2$ and $\alpha = 1$ it also accurately predicts $\bar{Y} = 14$.

**Proof of Proposition 4:**

An expression for $\partial \bar{Y}(0,0)/\partial\mu$ can be computed immediately from the derivative in Proposition 3. So we turn to the derivative of $\bar{Y}$ w.r.t. $r$ fixing $\mu = 0$. For this we use the power series expansion of the solution to the o.d.e. and write a system of two equations, value matching and smooth pasting, in two unknowns, $c_1$ and $\bar{Y}$. The first equation is given by value matching multiplied by $r$, where we have replaced the coefficient $c_0$ from its expression in terms of $c_1$ and where we have written the other coefficients, $c_m(c_1, r)$ as functions of $c_1$ and $r$. We have:

$$
0 = n(c_1\sigma^2 + \alpha) + r\Phi + r\left[ c_1\bar{Y} + \sum_{m=2}^{\infty} c_m(c_1, r)\,\bar{Y}^m \right]
$$

$$
0 = c_1 + \sum_{m=2}^{\infty} m\,c_m(c_1, r)\,\bar{Y}^{m-1}
$$

where:

$$
c_2(c_1, r) = \frac{\beta/2 + rc_1}{2\sigma^2(n+2)}\ , \quad c_{m+1}(c_1, r) = \left(\frac{r}{\sigma^2}\right)^{m-1}\frac{c_2(c_1, m)}{(2m+n)(m+1)} \quad \text{for}\ \ m \geq 2\ .
$$

We will totally differentiate this system, for which it is useful to note that:

$$
\begin{aligned}
\frac{\partial c_2(c_1, 0)}{\partial c_1} &= 0\ ,\ \ \frac{\partial c_2(c_1, 0)}{\partial r} = \frac{c_1}{2\sigma^2(n+2)} = -\frac{\alpha}{2\beta\sigma^2(n+2)} \\
\frac{\partial c_3(c_1, 0)}{\partial c_1} &= 0\ ,\ \ \frac{\partial c_3(c_1, 0)}{\partial r} = \frac{1}{\sigma^2}\frac{c_2}{(4+n)3} = \frac{1}{\sigma^2(4+n)3}\frac{\beta}{\sigma^2(n+2)4} \\
\frac{\partial c_{m+1}(c_1, 0)}{\partial c_1} &= \frac{\partial c_{m+1}(c_1, 0)}{\partial r} = c_m(c_1, 0) = 0\ \ \text{for all}\ m \geq 3\ .
\end{aligned}
$$

We have:

$$
\begin{aligned}
0 \;=\;& \Phi + \left[ c_1 \bar{Y} + \sum_{m=2}^{\infty} c_m(c_1, r)\, \bar{Y}^m \right] + \left( n\sigma^2 + r \sum_{m=2}^{\infty} \frac{\partial c_m(c_1, r)}{\partial c_1} \bar{Y}^m \right) \frac{\partial c_1}{\partial r} \\
&+\; r \left[ c_1 + \sum_{m=2}^{\infty} c_m(c_1, r) m\, \bar{Y}^{m-1} \right] \frac{\partial \bar{Y}}{\partial r} \\
0 \;=\;& \sum_{m=2}^{\infty} m \frac{\partial c_m(c_1, r)}{\partial r} \bar{Y}^{m-1} + \left( 1 + \sum_{m=2}^{\infty} m \frac{\partial c_m(c_1, r)}{\partial c_1} \bar{Y}^{m-1} \right) \frac{\partial c_1}{\partial r} \\
&+\; \left( \sum_{m=2}^{\infty} m\,(m-1)\, c_m(c_1, r)\, \bar{Y}^{m-2} \right) \frac{\partial \bar{Y}}{\partial r}
\end{aligned}
$$

Evaluating the first equation at $r = 0$ we obtain

$$
\frac{\partial c_1}{\partial r} \;=\; -\frac{\Phi + \left[ c_1 \bar{Y} + c_2 \bar{Y}^2 \right]}{n\sigma^2}\,.
$$

Evaluating the second equation at $r = 0$ we obtain

$$
\frac{\partial \bar{Y}}{\partial r} \;=\; -\sum_{m=2}^{3} m \frac{\partial c_m(c_1, r)}{\partial r} \frac{1}{2\,c_2} \bar{Y}^{m-1} - \frac{1}{2\,c_2} \frac{\partial c_1}{\partial r}
$$

replacing the value of $c_2$:

$$
\frac{\partial \bar{Y}}{\partial r} \;=\; -\frac{\partial c_2(c_1, r)}{\partial r} \frac{2}{2\,c_2} \bar{Y} - \frac{\partial c_3(c_1, r)}{\partial r} \frac{3}{2\,c_2} \bar{Y}^2 + \frac{1}{2\,c_2} \frac{\Phi + \left[ c_1 \bar{Y} + c_2 \bar{Y}^2 \right]}{n\sigma^2}
$$

replacing the expressions for $\partial c_2 / \partial r$ and $\partial c_3 / \partial r$:

$$
\begin{aligned}
\frac{\partial \bar{Y}}{\partial r} \;=\;& -\frac{c_1}{2\sigma^2(n+2)} \frac{2}{2\,c_2} \bar{Y} - \frac{1}{\sigma^2} \frac{c_2}{(4+n)3} \frac{3}{2\,c_2} \bar{Y}^2 + \frac{2(n+2)}{\beta} \left( \frac{\Phi + \left[ c_1 \bar{Y} + c_2 \bar{Y}^2 \right]}{n} \right) \\
\;=\;& -\frac{1}{\sigma^2(n+2)} \frac{c_1}{2\,c_2} \bar{Y} - \frac{1}{\sigma^2(4+n)2} \bar{Y}^2 + \frac{2(n+2)}{n\beta} \left( c_1 \bar{Y} + c_2 \bar{Y}^2 \right) + \frac{2(n+2)}{n\beta} \Phi
\end{aligned}
$$

using that $\bar{Y} = -c_1/(2c_2)$ and $c_1 = -\alpha/\sigma^2$:

$$
\begin{aligned}
\frac{\partial \bar{Y}}{\partial r} \;=\;& \frac{1}{\sigma^2(n+2)} \bar{Y}^2 - \frac{1}{\sigma^2(4+n)2} \bar{Y}^2 + \frac{2(n+2)}{n\beta} \bar{Y} \left( c_1 + c_2 \bar{Y} \right) + \frac{2(n+2)}{n\beta} \Phi \\
\;=\;& \left[ \frac{(4+n)2 - (n+2)}{(n+2)\sigma^2(4+n)2} + \frac{2(n+2)}{n\beta} c_2 \right] \bar{Y}^2 - \frac{2(n+2)}{n\beta} \frac{\alpha}{\sigma^2} \bar{Y} + \frac{2(n+2)}{n\beta} \Phi
\end{aligned}
$$

using that $\bar{Y} = (\alpha/\beta)2(n+2)$:

$$
\frac{\partial \bar{Y}}{\partial r} \;=\; \left[ \frac{6+n}{(n+2)\sigma^2(4+n)2} + \frac{2(n+2)}{n\beta} c_2 - \frac{1}{n\sigma^2} \right] \bar{Y}^2 + \frac{2(n+2)}{n\beta} \Phi
$$

replacing back $c_2$:

$$
\begin{aligned}
\frac{\partial \bar{Y}}{\partial r} &= \left[\frac{6+n}{(n+2)\sigma^2(4+n)2} + \frac{2(n+2)}{n\beta}\frac{\beta}{4\sigma^2(n+2)} - \frac{1}{n\sigma^2}\right]\bar{Y}^2 + \frac{2(n+2)}{n\beta}\Phi \\
&= \left[\frac{6+n}{(n+2)\sigma^2(4+n)2} + \frac{1}{2\,n\sigma^2} - \frac{1}{n\sigma^2}\right]\bar{Y}^2 + \frac{2(n+2)}{n\beta}\Phi \\
&= \left[\frac{6+n}{(n+2)\sigma^2(4+n)2} - \frac{1}{2\,n\sigma^2}\right]\bar{Y}^2 + \frac{2(n+2)}{n\beta}\Phi \\
&= \frac{1}{2\sigma^2}\left[\frac{6+n}{(n+2)(4+n)} - \frac{1}{n}\right]\bar{Y}^2 + \frac{2(n+2)}{n\beta}\Phi \\
&= \frac{1}{2\sigma^2}\left[\frac{(6+n)n - (n+2)(4+n)}{(n+2)(4+n)n}\right]\bar{Y}^2 + \frac{2(n+2)}{n\beta}\Phi \\
&= \frac{1}{2\sigma^2}\left[\frac{6n+n^2-4n-8-n^2-2n}{(n+2)(4+n)n}\right]\bar{Y}^2 + \frac{2(n+2)}{n\beta}\Phi \\
&= -\frac{4}{\sigma^2(n+2)(4+n)n}\bar{Y}^2 + \frac{2(n+2)}{n\beta}\Phi
\end{aligned}
$$

and replacing back $\bar{Y}$:

$$
\begin{aligned}
\frac{\partial \bar{Y}}{\partial r} &= -\frac{4(n+2)^2 4}{\sigma^2(n+2)(4+n)n}\left(\frac{\alpha}{\beta}\right)^2 + \frac{2(n+2)}{n\beta}\Phi \\
&= -\frac{16(n+2)}{(4+n)n}\frac{1}{\sigma^2}\left(\frac{\alpha}{\beta}\right)^2 + \frac{2(n+2)}{n}\frac{\Phi}{\beta} \\
&= \frac{2(n+2)}{n}\left[-\frac{8}{(4+n)}\frac{1}{\sigma^2}\left(\frac{\alpha}{\beta}\right)^2 + \frac{\Phi}{\beta}\right]
\end{aligned}
$$

Now we use the expression of the two derivatives in the expansion to obtain:

$$
\begin{aligned}
\bar{Y} &= \frac{\alpha}{\beta}2(n+2) + \frac{2(n+2)}{n}\left[-\frac{8}{(4+n)}\frac{r}{\sigma^2}\left(\frac{\alpha}{\beta}\right)^2 + \frac{r}{\beta}\Phi\right] \qquad\qquad\text{(A-12)}\\
&+ 16\left(\frac{\alpha}{\beta}\right)^2\frac{2+n}{4+n}\frac{\mu}{\sigma^2} + o\left(||(\mu,r)||\right)
\end{aligned}
$$

which after rearranging gives the desired expression. $\square$

# I  Derivative of $\overline{Y}$ with respect to $\sigma^2$ at $r = (n-1)\,\mu$

In this appendix we show that when $r = (n-1)\,\mu$ then $\overline{Y}$ is independent of $\sigma$.

Recall that the power series solution (11) is

$$
V(Y) = \sum_{m=0}^{\infty} c_m\, Y^m,
$$

46

where $c_0$ is to be determined, and, defining $\tau = 1/\sigma^2$, we have recursively

$$c_1 = \frac{(r\,c_0 - n\,\alpha)\,\tau}{n - 1},$$

$$c_2 = \frac{\left[\frac{1}{2}\,\beta + (2\,\mu + r)\,c_1\right]\,\tau}{2\,(n + 1)},$$

$$c_{m+1} = \frac{(2\,\mu\,m + r)\,\tau}{(2\,m + n - 1)\,(m + 1)}\,c_m \quad \text{for all } m \geq 2.$$

For $m = 0, 1, \ldots$ define

$$f_m = \frac{2\,\mu\,m + r}{(2\,m + n - 1)\,(m + 1)}, \tag{A-13}$$

and

$$F_0 = 1, \quad F_{m+1} = f_m\,F_m = \prod_{k=0}^{m} f_k. \tag{A-14}$$

to write the equations for the coefficients as

$$c_1 = f_0\,c_0\,\tau - \frac{n}{n - 1}\,\alpha\,\tau, \tag{A-15}$$

$$c_2 = \left[\frac{\beta}{4\,(n + 1)} + f_1\,c_1\right]\,\tau, \tag{A-16}$$

$$c_m = \frac{F_m}{F_2}\,\tau^{m-2}\,c_2 \quad \text{for all } m \geq 3. \tag{A-17}$$

In the power series solution, we can take any one of the coefficients $c_m$ as an independent variable, and use (A-15), (A-16) and (A-17) to express all the others in terms of that one. It will prove most convenient to choose $c_2$ for this role. Write the solution as $V(Y; c_2, \theta)$ where $\theta$ will stand for any of the parameters $r$, $\alpha$, $\beta$, $\mu$ or $\tau$, and the $c_m$ for $m = 0, 1, 3, \ldots$ are all expressed as functions of $c_2$ and $\theta$. Then $\overline{Y}$ and $c_2$ are defined by the value matching and smooth pasting conditions

$$V(\overline{Y}; c_2, \theta) = -\Phi \tag{A-18}$$

$$V'(\overline{Y}; c_2, \theta) = 0 \tag{A-19}$$

Taking total differentials,

$$V'\,d\overline{Y} + \frac{\partial V}{\partial c_2}\,dc_2 + \frac{\partial V}{\partial \theta}\,d\theta = 0,$$

$$V''\,d\overline{Y} + \frac{\partial V'}{\partial c_2}\,dc_2 + \frac{\partial V'}{\partial \theta}\,d\theta = 0,$$

47

where the various derivatives of $V$ are understood to be evaluated at $(\overline{Y}; c_2, \theta)$. Using (A-19) in the first of these and substituting into the second,

$$V'' \, d\overline{Y} = \left\{ \frac{\partial V'/\partial c_2}{\partial V/\partial c_2} \frac{\partial V}{\partial \theta} - \frac{\partial V'}{\partial \theta} \right\} d\theta \, . \qquad \text{(A-20)}$$

Since $V''(\overline{Y}; c_2, \theta) > 0$, the sign of $d\overline{Y}/d\theta$ is same as the sign of the expression in the large brackets on the right hand side.

From (A-17) we have, for all $m \geq 3$,

$$\frac{\partial c_m}{\partial c_2} = \frac{F_m}{F_2} \tau^{m-2} \, , \qquad \text{(A-21)}$$

and, differentiating (A-15) and (A-16) using (A-13) and (A-14), it is easy to verify that the same extends to $m = 0$ and $1$ also.

Therefore

$$
\begin{aligned}
\frac{\partial V}{\partial c_2} &= \sum_{m=0}^{\infty} \frac{\partial c_m}{\partial c_2} Y^m \\
&= \sum_{m=0}^{\infty} \frac{F_m}{F_2} \tau^{m-2} Y^m \\
&= \frac{\tau^{-2}}{F_2} \sum_{m=0}^{\infty} F_m \tau^m Y^m
\end{aligned}
\qquad \text{(A-22)}
$$

and

$$V' = \sum_{m=0}^{\infty} m \, c_m Y^{m-1} \, ,$$

(note: including the $m = 0$ term is harmless) therefore

$$
\begin{aligned}
\frac{\partial V'}{\partial c_2} &= \sum_{m=0}^{\infty} m \frac{\partial c_m}{\partial c_2} Y^{m-1} \\
&= \sum_{m=0}^{\infty} \frac{F_m}{F_2} \tau^{m-2} m Y^{m-1} \\
&= \frac{\tau^{-2}}{F_2} \sum_{m=0}^{\infty} m F_m \tau^m Y^{m-1}
\end{aligned}
\qquad \text{(A-23)}
$$

Observe that $\partial V'/\partial c_2 > 0$; therefore from (A-20) the sign of $d\overline{Y}/d\theta$ is the same as that of

$$\frac{\partial V/\partial \theta}{\partial V/\partial c_2} - \frac{\partial V'/\partial \theta}{\partial V'/\partial c_2} \qquad \text{(A-24)}$$

The key parameter of concern for us is $\tau \equiv \sigma^{-2}$. Since $c_2$ is the chosen independent variable, from (A-17) we have

$$\frac{\partial c_m}{\partial \tau} = (m-2) \, \frac{F_m}{F_2} \, \tau^{m-3} \, c_2 \quad \text{for all } m \geq 3\,, \tag{A-25}$$

which can be harmlessly extended to $m = 2$. Next, from (A-16),

$$c_1 = \frac{c_2}{f_1} \, \tau^{-1} - \frac{\beta}{4(n+1) \, f_1} = \frac{F_1}{F_2} \, c_2 \, \tau^{-1} - \frac{F_1}{F_2} \, \frac{\beta}{4(n+1)}\,, \tag{A-26}$$

and

$$\begin{aligned}
c_0 &= \frac{1}{f_0} \, \tau^{-1} \, c_1 + \frac{\alpha}{f_0} \, \frac{n-1}{n} \\
&= \frac{1}{f_0} \, \tau^{-1} \left[ \frac{F_1}{F_2} \, c_2 \, \tau^{-1} - \frac{F_1}{F_2} \, \frac{\beta}{4(n+1)} \right] + \frac{\alpha}{f_0} \, \frac{n-1}{n} \\
&= \frac{\alpha}{f_0} \, \frac{n-1}{n} + \frac{F_0}{F_2} \, \tau^{-2} \, c_2 - \frac{F_0}{F_2} \, \frac{\beta}{4(n+1)} \, \tau^{-1}
\end{aligned} \tag{A-27}$$

Therefore

$$\begin{aligned}
\frac{\partial V}{\partial \tau} &= \frac{\partial c_0}{\partial \tau} + \frac{\partial c_1}{\partial \tau} \, Y + \sum_{m=3}^{\infty} \frac{\partial c_m}{\partial \tau} \, Y^m \\
&= -2 \frac{F_0}{F_2} \, \tau^{-3} \, c_2 + \frac{F_0}{F_2} \, \frac{\beta}{4(n+1)} \, \tau^{-2} - \frac{F_1}{F_2} \, c_2 \, \tau^{-2} \, Y + \sum_{m=3}^{\infty} (m-2) \, \frac{F_m}{F_2} \, \tau^{m-3} \, c_2 \, Y^m \\
&= \frac{F_0}{F_2} \, \frac{\beta}{4(n+1)} \, \tau^{-2} + \frac{c_2}{F_2} \, \tau^{-3} \sum_{m=0}^{\infty} (m-2) \, F_m \, \tau^m \, Y^m\,.
\end{aligned} \tag{A-28}$$

Observe how some terms for $m = 0$ and $1$ have been absorbed into the general sum, and a term for $m = 2$ has been harmlessly added. Finally,

$$\frac{\partial V'}{\partial \tau} \equiv \frac{\partial}{\partial Y} \left( \frac{\partial V}{\partial \tau} \right) = \frac{c_2}{F_2} \, \tau^{-3} \sum_{m=0}^{\infty} m \, (m-2) \, F_m \, \tau^m \, Y^{m-1} \tag{A-29}$$

So the condition for $\partial \overline{Y} / \partial \sigma < 0$ or $\partial \overline{Y} / \partial \tau > 0$ is

$$\frac{\frac{F_0}{F_2} \frac{\beta}{4(n+1)} \tau^{-2} + \frac{c_2}{F_2} \tau^{-3} \sum_{m=0}^{\infty} (m-2) F_m \, \tau^m \, \overline{Y}^m}{\frac{\tau^{-2}}{F_2} \sum_{m=0}^{\infty} F_m \, \tau^m \, \overline{Y}^m} > \frac{\frac{c_2}{F_2} \tau^{-3} \sum_{m=0}^{\infty} m \, (m-2) \, F_m \, \tau^m \, \overline{Y}^{m-1}}{\frac{\tau^{-2}}{F_2} \sum_{m=0}^{\infty} m \, F_m \, \tau^m \, \overline{Y}^{m-1}}\,.$$

Canceling some common factors and using $F_0 = 1$, this becomes

$$\frac{\frac{\tau}{c_2} \frac{\beta}{4(n+1)} + \sum_{m=0}^{\infty} (m-2) F_m \, \tau^m \, \overline{Y}^m}{\sum_{m=0}^{\infty} F_m \, \tau^m \, \overline{Y}^m} > \frac{\sum_{m=0}^{\infty} m \, (m-2) \, F_m \, \tau^m \, \overline{Y}^{m-1}}{\sum_{m=0}^{\infty} m \, F_m \, \tau^m \, \overline{Y}^{m-1}}\,. \tag{A-30}$$

49

We want to prove that when $r = (n-1)\,\mu$, this holds as an equality. In this situation,

$$f_m = \frac{2\,\mu\,m + (n-1)\,\mu}{(2\,m + n - 1)(m+1)} = \frac{\mu}{m+1}\,. \tag{A-31}$$

Therefore, using (A-14),

$$F_{m+1} = \prod_{k=0}^{m} \frac{\mu}{k+1} = \frac{\mu^{m+1}}{(m+1)!}\,. \tag{A-32}$$

This simplifies the various expressions in the inequality (A-30). We have

$$\sum_{m=1}^{\infty} F_m\,\tau^m\,\overline{Y}^m = \sum_{m=0}^{\infty} \frac{1}{m!}\,(\mu\,\tau\,\overline{Y})^m = e^{\mu\,\tau\,\overline{Y}} = e^J,$$

introducing, for convenience of notation, the abbreviation $J = \mu\,\tau\,\overline{Y}$.

Next

$$\begin{aligned}
\sum_{m=0}^{\infty} m\,F_m\,\tau^m\,\overline{Y}^m &= \sum_{m=1}^{\infty} m\,\frac{\mu^m}{m!}\,\tau^m\,\overline{Y}^m \\
&= \mu\,\tau\,\overline{Y} \sum_{m=1}^{\infty} \frac{1}{(m-1)!}\,\mu^{m-1}\,\tau^{m-1}\,\overline{Y}^{m-1} \\
&= \mu\,\tau\,\overline{Y} \sum_{m=0}^{\infty} \frac{1}{m!}\,J^m = J\,e^J
\end{aligned}$$

Finally

$$\begin{aligned}
\sum_{m=0}^{\infty} m\,(m-2)\,F_m\,\tau^m\,\overline{Y}^m &= \sum_{m=1}^{\infty} \frac{m-2}{(m-1)!}\,\mu^m\,\tau^m\,\overline{Y}^m \\
&= J \sum_{m=1}^{\infty} \frac{m-2}{(m-1)!}\,J^{m-1} = J \sum_{m=0}^{\infty} \frac{m-1}{m!}\,J^m \\
&= J \sum_{m=0}^{\infty} \frac{m}{m!}\,J^m - J \sum_{m=0}^{\infty} \frac{1}{m!}\,J^m \\
&= J^2 \sum_{m=1}^{\infty} \frac{1}{(m-1)!}\,J^{m-1} - J\,e^J \\
&= J^2 \sum_{m=0}^{\infty} \frac{1}{m!}\,J^m - J\,e^J \\
&= J^2\,e^J - J\,e^J = J\,(J-1)\,e^J
\end{aligned}$$

Using all these, the (A-30) is

$$\frac{\frac{\tau}{c_2}\,\frac{\beta}{4(n+2)} + J\,e^J - 2\,e^J}{e^J} > \frac{J\,(J-1)\,e^J\,/\,Y}{J\,e^J\,/\,Y}\,,$$

50

or

$$\frac{\tau}{c_2} \frac{\beta}{4(n+2)} + J\, e^J - 2\, e^J > (J-1)\, e^J,$$

or

$$\frac{\tau}{c_2} \frac{\beta}{4(n+2)} > e^J. \tag{A-33}$$

Now consider the smooth pasting condition

$$
\begin{aligned}
0 = V'(\overline{Y}) &= c_1 + \sum_{m=2}^{\infty} m\, c_m\, \overline{Y}^{m-1} \\[2mm]
&= \frac{1}{f_1}\left[ c_2\, \tau^{-1} - \frac{\beta}{4(n+1)} \right] + \sum_{m=2}^{\infty} m\, \frac{F_m}{F_2}\, \tau^{m-2}\, c_2\, \overline{Y}^{m-1} \quad \text{using (A-16) and (A-17)} \\[2mm]
&= -\frac{F_1}{F_2} \frac{\beta}{4(n+1)} + \frac{c_2}{F_2} \sum_{m=1}^{\infty} m\, F_m\, \tau^{m-2}\, \overline{Y}^{m-1} \quad \text{using (A-14) for } m=1 \\[2mm]
&= -\frac{F_1}{F_2} \frac{\beta}{4(n+1)} + \frac{c_2}{F_2} \sum_{m=1}^{\infty} m\, \frac{\mu^m}{m!}\, \tau^{m-2}\, \overline{Y}^{m-1} \\[2mm]
&= -\frac{F_1}{F_2} \frac{\beta}{4(n+1)} + \frac{c_2}{F_2} \sum_{m=1}^{\infty} \frac{1}{(m-1)!}\, \mu^m\, \tau^{m-2}\, \overline{Y}^{m-1} \\[2mm]
&= -\frac{F_1}{F_2} \frac{\beta}{4(n+1)} + \frac{c_2}{F_2}\, \mu\, \tau^{-1} \sum_{m=1}^{\infty} \frac{1}{(m-1)!}\, (\mu\, \tau\, \overline{Y})^{m-1} \\[2mm]
&= -\frac{F_1}{F_2} \frac{\beta}{4(n+1)} + \frac{c_2}{F_2}\, \mu\, \tau^{-1}\, e^J
\end{aligned}
$$

Therefore

$$\frac{\tau\, F_1}{\mu\, c_2} \frac{\beta}{4(n+1)} = e^J$$

or

$$\frac{\tau}{c_2} \frac{\beta}{4(n+2)} = e^J \quad \text{using (A-32) for } m = 0.$$

Thus (A-33) holds as an equality.

## J  Discrete time *i.i.d.* case

Consider the case where $\mu$ goes to infinity, so $Y$ is iid. In particular consider the discrete time model with

$$X_i(t+\Delta) = X_i(t) - \mu_\Delta \Delta X_i(t) + \sqrt{\Delta}\sigma_\Delta W_i + \sqrt{\Delta}\sigma_{c,\Delta} W_c \tag{A-34}$$

for a period of length $\Delta$ where we index the drift and volatility by $\Delta$ and were $W_i, W_c$ are iid normal random variables. Set $\mu_\Delta\, \Delta = 1$ and thus we have

$$X_i(t+\Delta) = \sqrt{\Delta}\, \sigma_\Delta\, W_i + \sqrt{\Delta}\, \sigma_{c,\Delta}\, W_c \tag{A-35}$$

and taking into account the union-wide policy $Z$:

$$x_i(t + \Delta) \equiv X_i(t + \Delta) - Z(t + \Delta) = \sqrt{\Delta}\, \sigma_\Delta \left( W_i - \frac{1}{n} \sum_{j=1}^{n} W_j \right)$$

thus the collective state variable is given by

$$Y \equiv \sum_{i=1}^{n} [x_i(t + \Delta)]^2 = \Delta\, \sigma_\Delta^2 \sum_{i=1}^{n} \left( \frac{n-1}{n} W_i - \frac{1}{n} \sum_{j \neq i}^{n} W_j \right)^2 \tag{A-36}$$

Let $\bar{\sigma}^2$ the unconditional variance of the continuous time process, which satisfies: $\bar{\sigma}^2 = \sigma^2/(2\mu)$. We can also let $\sigma_\Delta$ change with $\Delta$ and $\mu_\Delta$ so that the unconditional variance stays constant with $\Delta$, which gives

$$\Delta\, \sigma_\Delta^2 = \bar{\sigma}^2\, 2\, \mu_\Delta\, \Delta = 2\, \bar{\sigma}^2 \tag{A-37}$$

and thus

$$Y \equiv \sum_{i=1}^{n} [x_i(t + \Delta)]^2 = 2\, \bar{\sigma}^2 \sum_{i=1}^{n} \left( \frac{n-1}{n} W_i - \frac{1}{n} \sum_{j \neq i}^{n} W_j \right)^2 \tag{A-38}$$

In this case we will make the comparative static with respect to $\bar{\sigma}^2$.

Thus, when $X_i$ are iid and the time period is of length $\Delta > 0$ the value function becomes:

$$V(Y\,;\bar{\sigma}^2) = \tag{A-39}$$
$$\max \left\{ -\Phi\,, \ \Delta(\alpha n - \frac{\beta}{2} Y) + \frac{1}{1 + r\Delta} E \left[ V \left( 2\, \bar{\sigma}^2 \sum_{i=1}^{n} \left( \frac{n-1}{n} W_i - \frac{1}{n} \sum_{j \neq i}^{n} W_j \right)^2\,; \bar{\sigma}^2 \right) \right] \right\}$$

The threshold $\overline{Y}$ solves:

$$\overline{Y} = \frac{\alpha n}{\beta/2} + \frac{\Phi}{\Delta\beta/2} + \frac{1}{(1 + r\Delta)\beta/2} E \left[ V \left( 2\, \bar{\sigma}^2 \sum_{i=1}^{n} \left( \frac{n-1}{n} W_i - \frac{1}{n} \sum_{j \neq i}^{n} W_j \right)^2\,; \bar{\sigma}^2 \right) \right] \tag{A-40}$$

Note that for any $\Delta > 0$ the operator defined by the left hand side of equation (A-39) is a contraction. First we argue that for a given $Y$, the function $V(Y\,;\bar{\sigma}^2)$ is decreasing in $\bar{\sigma}^2$. This can be shown by a guess and verify argument. Second we argue that the value function is decreasing in $\bar{\sigma}^2$, which can also be shown using the previous result and a guess and verify strategy. Finally, using these two results it is immediate to show that the optimal threshold is also decreasing in $\bar{\sigma}^2$.

# K   Baseline Parameter Values

In this appendix we discuss the literature which we use to select our benchmark parameter values.

**Parameters $\sigma$ and $\mu$.**

Engel and Rogers (2001) reports standard deviation for 12-month changes in CPI measured in the same currency across cities in Europe located in different countries of 0.072 for the period 1981-1997.[11]

Crucini and Telmer (2012) compute the variance of 12-month price changes for 300 comparable goods and services for 123 cities and 78 countries for the period 1990-2005. They estimate that the common component of the variance of log price changes for cities located in different countries is 0.01, which gives a 0.1 standard deviation.[12]

A standard reference for estimates of variability on changes on real exchange rates, measured using CPI is Mussa (1986); this gives estimates of this standard deviation above 0.07 for major european countries. He argues, convincingly, that real exchange rates tend to be more volatile when nominal exchange rates are floating. [13]

The surveys Rogoff (1996) and Taylor and Taylori (2004) report point estimates of the half-life of (relative) PPP deviation typically in the 3-5 years range. Yet these estimates are quite imprecise, since reliable estimation of the speed of convergence requires either very long data sets or the pooling of the experience of many countries, see Murray and Papell (2002) and Rossi (2005).

---

[11]See Table 2, column variance of $\Delta P(j,k)$ for 12-months, row International, which gives a variance of $100\times$ log changes of 52.3. This is the mean variance of all pairs of cities across located in different countries. This variance is much larger than the one intra-nationally, which is 0.96. For comparison Engel and Rogers (1996) report that the average standard deviation across 14 goods of the two months price changes in common currency across US and Canadian cities to be 0.0367 for the period 1978 to 1994. Annualizing this standard deviation gives 0.0899.

[12]See Table 4 "Variance of Changes in LOP Deviations", rows OECD countries, International, column across goods (common), which gives a variance of 0.010. This is much larger than the variance across the same countries, which is 0.001

[13]Mussa (1986) computed variance of (log) changes on quarterly real exchange rates, where each quarterly real exchange rate is computed by averaging nominal exchange rates during the quarter and dividing it by CPI. The figures are for 1957-1984 period. For the sub-period where nominal exchange rates are pegged, he finds that the variance is about half of the variance for the whole period. To annualize this quarterly variance we multiplied it by 4. Due to the assumed mean reversion, and the averaging used in this construction, this volatility are lower bounds for $\sigma$.

Engel and Rogers (2001) pooled recent data from many countries and find no evidence of different degree of mean reversion for PPP deviation in currency unions. Lothian and Taylor (1996) use long time series argue for similar speed of mean reversion on floats and pegs.

Cecchetti, Mark, and Sonora (2002) use long time series and CPI for US states to find that relative PPP deviations have a half-life of about 9 years.

Using prices for comparable goods across cities around the world, Rogers (2002) finds a decrease in the cross-sectional dispersion of prices across countries in the Euro area to levels similar to the ones of US states, which had happened mostly before the establishment of the euro. Using the Eurostat PPP price index, Lane (2006) finds that while the annual dispersion in inflation rates have not been much different to the variation across US regions, inflation differentials in the euro area have been more persistent.

**Parameters $\alpha$ and $\Phi$.**

Mendizabal (2002) estimates reduction on transaction cost as high as 0.69% of GDP for members countries, and EEC (1990) estimated them to be about 0.4% of GDP. The other source are gains from trade. We use the lower bound of Rose. (2008) estimates, which gives a long run increase of trade of 30%. We translate the predicted increase in trade as it would have been due to a reduction on trade cost, and write its welfare-equivalent units of GDP using the method of Rodriguez-Clare, Arkolakis, and Costinot (2012), which gives a value of about 2% of GDP. [14] To be on the conservative side we select a value of $\alpha$ of 0.02.

Sandleris and Wright (2011) studied in detail the cost of the 2001 argentinean crises, for which they find a cost of about 25% of GDP. This case of which seems appropriate given that dollars were a legal tender before the public debt default and conversion of asset and liabilities from dollars to pesos. Laeven and Valencia (2008) compile an extensive database with description of statistics around crises and a brief narrative. The experience of argentina is among the costlier, as measured as GDP performance around the crisis time, but by no means the most dramatic one. We set the value of $\Phi = n\ 0.20$. These high values associated with financial distressed are not inconsistent with the ones estimated in the corporate finance literature as the economic cost associated with financial distressed firms.[15]

---

[14]We use an import to GDP ratio of 30% and an Armington elasticity $\epsilon$ of 7, where ratio of the real income is $W'/W = (\lambda'/\lambda)^{-1/\epsilon}$ where $\lambda'$ and $\lambda$ are the domestic absorption after the currency union and before respectively. Thus we have $W'/W = ((1 - 0.3 \times 1.3)/(1 - 0.3))^{-1/7} = 1.0199$ i.e. an approximately 2% gain.

[15]Weiss (1990) estimates the direct cost of reorganization to be 3% of total value of the firm. Andrade and Kaplan (1998) and Davydenko, A., and Xiaofei (2012) estimates the direct and indirect economic cost

**Parameter $\beta$.**

First consider an economy with two tradable goods, and a symmetric CES utility function

$$U(q_1, q_2) = \left[ q_1^{1-1/\eta} + q_2^{1-1/\eta} \right]^{\eta/(\eta-1)}$$

where consumers face a price of the fist good $e^x$ and the second has a price normalized to one. We interpret good 1 to be the local good, and the good 2 to be the foreign good. Production must satisfy $q_1 + q_2 = 1$. We let $V(x)$ the indirect utility of the representative consumer at an equilibrium where the relative price of the first good is $x$. We have

$$V(x) = U\left(q(x) \ , \ 1 - q(x)\right) \quad \text{where} \quad q(x) \ \text{solves} \ \frac{U_1\left(q(x) \ , \ 1 - q(x)\right)}{U_2\left(q(x) \ , \ 1 - q(x)\right)} = e^x \ .$$

We have $\frac{q(x)}{1-q(x)} = e^{-\eta x}$ or $q(x) = \frac{e^{-\eta x}}{1+e^{-\eta x}}$ so $q(0) = 1/2$ and $q'(0) = -\eta/2$. Differentiating $V$ we have

$$V'(x) = \left[U_1(q, 1-q) - U_2(q, 1-q)\right] q'(x)$$

and

$$V''(x) = \left[U_1(q, 1-q) - U_2(q, 1-q)\right] q''(x) + \left[U_{11}(q, 1-q) - 2U_{12}(q, 1-q) + U_{22}(q, 1-q)\right] q'(x)^2$$

Expanding $V$ around $x = 0$ and ignoring terms cubic and higher:

$$
\begin{aligned}
V(x) &\approx U(q(0), 1 - q(0)) + \frac{1}{2}\left[U_{11}(q(0), 1 - q(0)) - 2U_{12}(q(0), 1 - q(0)) + U_{22}(q(0), 1 - q(0))\right] q'(0)^2 x^2 \\
&= U(q(0), 1 - q(0)) + \left[U_{11}(q(0), 1 - q(0)) - U_{12}(q(0), 1 - q(0))\right] q'(0)^2 x^2
\end{aligned}
$$

where we use that $U_1(q(0), 1 - q(0)) = U_2(q(0), 1 - q(0))$ and symmetry. Computing the derivatives we have:

$$
\begin{aligned}
U_1 &= \left[q_1^{1-1/\eta} + q_2^{1-1/\eta}\right]^{\eta/(\eta-1)-1} q_1^{-1/\eta} \\
U_{12} &= \left(1 - \frac{1}{\eta}\right)\left[q_1^{1-1/\eta} + q_2^{1-1/\eta}\right]^{\eta/(\eta-1)-2} \left(\eta/(\eta-1) - 1\right) q_1^{-1/\eta} q_2^{-1/\eta} \\
U_{11} &= \left(1 - \frac{1}{\eta}\right)\left[q_1^{1-1/\eta} + q_2^{1-1/\eta}\right]^{\eta/(\eta-1)-2} \left(\eta/(\eta-1) - 1\right) q_1^{-1/\eta} q_1^{-1/\eta} \\
&\quad - \frac{1}{\eta}\left[q_1^{1-1/\eta} + q_2^{1-1/\eta}\right]^{\eta/(\eta-1)-1} q_1^{-1/\eta-1}
\end{aligned}
$$

---

of financial distress to be in the order of 10% and 20% (resp.) of the total value of the firm. Thus if the flow value of a firm is 5 times its value -a very low multiple for a corporation, but a more appropriate value for a whole economy-, and taking a 15% to be the cost of re-organization, then if 30% of the asset of the economy are in firms that need to be reorganized, one will obtain a cost of approximately 20%, as in our benchmark value.

Thus, dividing by $U(q(0), 1 - q(0))$, to compute the change in equivalent consumption units:

$$
\frac{V(x)}{U(q(0), 1 - q(0))} \approx 1 - \frac{1}{\eta} \frac{\left[q_1^{1-1/\eta} + q_2^{1-1/\eta}\right]^{\eta/(\eta-1)-1}}{\left[q_1^{1-1/\eta} + q_2^{1-1/\eta}\right]^{\eta/(\eta-1)}} q_1^{-1/\eta-1} \left(\frac{\eta}{q(0)}\right)^2 x^2
$$

$$
= 1 - \frac{1}{\eta} \frac{q(0)^{-1/\eta-1}}{[q(0)^{1-1/\eta} + q(0)^{1-1/\eta}]} \left(\frac{\eta}{q(0)}\right)^2 x^2
$$

$$
= 1 - \frac{1}{\eta} \frac{q(0)^{1-1/\eta}}{[q(0)^{1-1/\eta} + q(0)^{1-1/\eta}]} q(0)^2 \left(\frac{\eta}{q(0)}\right)^2 x^2
$$

$$
= 1 - \frac{q(0)^{1-1/\eta}}{[q(0)^{1-1/\eta} + q(0)^{1-1/\eta}]} \eta \, x^2 = 1 - \frac{1}{2} \eta \, x^2
$$

Now consider an economy with a non-tradable sector, which preference $\hat{U}(q_N, q_T)$ where $q_T = U(q_1, q_2)$. In the case where resources are not substitutable across the tradable and non-tradable sectors, the change in welfare is given by the change in utility of the tradable goods times its share. In this case we will obtain the desired expression.

# L   Optimality of one country's exit

Using the law of motion of $x_i = X - z$, the definition of $y_i$, and Ito's lemma we obtain:

$$
\begin{aligned}
dy_i &= 2 x_i \, dx + \frac{1}{2} 2 \sigma^2 \, \mathrm{E} \left( \frac{n-1}{n} dw_i - \frac{1}{n} \sum_{j \neq i} dw_j \right)^2 = 2 x_i \, dx + \frac{1}{2} 2 \sigma^2 \frac{n-1}{n} dt \\
&= \left[ \sigma^2 \frac{n-1}{n} - 2 \mu x_i^2 \right] dt + 2 x_i \sigma \left( \frac{n-1}{n} dw_i - \frac{1}{n} \sum_{j \neq i}^{n} dw_j \right) \\
&= \left[ \sigma^2 \frac{n-1}{n} - 2 \mu y_i \right] dt + 2 \sigma \sqrt{\frac{n-1}{n}} y_i \, dw_y \quad\quad\quad\quad\quad \text{(A-41)}
\end{aligned}
$$

Since

$$
dY = \left[(n-1)\sigma^2 - 2\mu Y\right] dt + 2\sigma \sum_{i=1}^{n} x_i \, dw_i = \left[(n-1)\sigma^2 - 2\mu Y\right] dt + 2\,\sigma \sqrt{Y}\, dw
$$

Then

$$
\begin{aligned}
\mathrm{E}\left[ dy_i \, dY \right] &= \mathrm{E} \left[ 2 x_i \sigma \left( \frac{n-1}{n} dw_i - \frac{1}{n} \sum_{j \neq i}^{n} dw_j \right) \left( 2\sigma \sum_{s=1}^{n} x_s \, dw_s \right) \right] \\
&= 4\,\sigma^2 x_i^2 \, dt - 4\,\sigma^2 \frac{x_i}{n} \left( x_i + \sum_{s \neq i}^{n} x_s \right) dt = 4\,\sigma^2 y_i \, dt
\end{aligned}
$$

Note that since $\sum_{i=1}^{n} x_i = 0$ then for a given $Y$, let $y_i = x$ and let $y_{-i} = -mx$ for $1 \leq m \leq n - 1$ then $Y = y + \frac{m}{m^2} y$ so $Y = y \frac{m}{m+1}$ then setting $m = n - 1$, its highest value, we obtain $y \leq \frac{n-1}{n} Y$. Letting $w_y$ and $w$ the BM's of the process for $y$ and $Y$ respectively we have that:

$$4 \, \sigma^2 \, y \, dt = \mathrm{E} \, [\, dy \, dY \,] = 4 \, \sigma^2 \, \sqrt{y \, \frac{n-1}{n}} \, Y \; \mathrm{E} \, [\, dw_y \, dw \,] \;,$$

and thus

$$\mathrm{E} \, [\, dw_y \, dw \,] = \sqrt{\frac{y}{Y} \frac{n}{n-1}} \, dt \; .$$

To show that $\min_{y,Y} v \, (Y \,, y) \leq -\Phi/n$ note that $v(Y, Y\frac{n-1}{n}) < v(Y, Y\frac{1}{n})$ for all $Y < \overline{Y}$, which follows because the flow return function used to construct $v(Y, y)$ is strictly decreasing in $y$, and because the Markov process for $y$ is monotone. Then since $v \left( Y \,, \frac{Y}{n} \right) = V(Y)/n$ for all $0 \leq Y \leq \overline{Y}$, and $V(\overline{Y}) = -\Phi/n$. $\square$

# M  Finite Difference Approximation to $v$

Here we describe the discretization procedure we follow to compute $v(y, Y)$, following Kushner and Dupuis (2001). First, it is convenient to rescale $y$ as $g = \delta \, y$. We then have:

$$dg = \left[ \delta \, \sigma^2 \frac{n-1}{n} - 2 \, \mu \, g \right] dt + 2 \sqrt{\delta \sigma^2 \, g \, \frac{n-1}{n}} \, dw_y \tag{A-42}$$

with $\mathbb{E} \, [dg \, dY] = 4 \, \sigma^2 \, g \, dt$ and $\mathbb{E} \, [dg^2] = 4 \, \sigma^2 \, g \, \delta \, \frac{n-1}{n}$ . We consider the following discrete time, discrete state Markov process approximation around a point $w \equiv (Y, g)$. At each point $w$ for which $w = (Y, g)$ and $Y < \overline{Y}$.

We will set

$$\delta = n/(n - 1) \text{ so that } W \equiv \{w = (g, Y) \in \mathbb{R}_+^2 : 0 \leq g \leq Y \text{ and } 0 \leq Y \leq \overline{Y}\}$$

We will consider a grid on $W$ and use the formulae described above for the probabilities for all values for which $0 < g < Y$ and $0 < Y$ and $Y < \overline{Y}$. On the boundary of the set $W$ we use different expressions since this case is not explicitly considered in Kushner and Dupuis (2001)[16]

We use $b(w)$ is a two dimensional vector with the drift of $Y$ and $g$ given our choice of $\delta$:

$$b_1(w) = \sigma^2 \, (n - 1) \; - \; 2\mu \, Y \text{ and } b_2(w) = \delta \, \sigma^2 \frac{n-1}{n} - 2 \, \mu \, g$$

---

[16]There are two features that makes the diffusion degenerate: the first is that as $Y$ goes to zero the system becomes deterministic, and the second is that for $g = Y$ the two processes are perfectly correlated. Additionally, we want to design a discrete time discrete state Markov process approximation that keeps the state on the feasible set for the continuous time solution.

and $a_{ij}(w)$ for the elements of the matrix:

$$a_{11}(w) \equiv \frac{1}{dt}\mathbb{E}[dY^2] = 4\,\sigma^2\,Y \ , \ a_{22}(w) \equiv \frac{1}{dt}\mathbb{E}[dg^2] = 4\,\sigma^2\,\delta\,g\,\frac{n-1}{n}\ \text{and}$$

$$a_{12}(w) \equiv \frac{1}{dt}\mathbb{E}[dg\,dY] = 4\,\sigma^2\,g \ .$$

We use $e_i$ is the unit vector in $\mathbb{R}^2$ with a one in position $i$. We are now ready to describe the expressions for each case:

(I) Case $g < Y < \overline{Y}$. In this case the process can move to any of the following eight adjacent position in the rectangular grid. The probabilities are given by:

$$\Pr\left(w \pm e_i h \mid w\right) \ = \ \left[a_{ii}(w)/2 - a_{ij}(w)/2 + hb_i^{\pm}(w)\right]\,/Q(w)$$
$$\Pr\left(w + e_i h + e_j h \mid w\right) \ = \ a_{ij}^{+}(w)\,/2Q(w)$$
$$\Pr\left(w - e_i h + e_j h \mid w\right) \ = \ a_{ij}^{-}(w)\,/2Q(w)$$

for $i, j = 1, 2$ and $i \neq j$, , where where the super-index $\pm$ denotes the positive and negative part of a scalar as in

$$d^{+} \ = \ |d| \text{ if } d > 0 \text{ and } = 0 \text{ otherwise, and}$$
$$d^{-} \ = \ |d| \text{ if } d < 0 \text{ and } = 0 \text{ otherwise},$$

where the factor $Q(w)$ is defined as

$$Q(w) \equiv a_{11}(w) + a_{22}(w) - a_{12}(w) + h|b_1(w)| + |b_2(w)|$$

Notice that since $\delta = n/(n-1)$ we have $a_{22}(w) \geq a_{12}(w)$, and thus all the probabilities are non-negative. The length of the time period at node $w$ is given by

$$\Delta t(w) = \frac{h^2}{Q(w)}$$

Note that if $w = (Y, 0)$ or then $g' \geq g$ with probability one, and if $w = (0, 0)$ then both $g'$ and $Y'$ are weakly higher with probability one.

(II) Case $g = Y$ or $w = (g, g)$ and $0 \leq Y = g < \overline{Y}$. Note that $a_{11}(w) = a_{22}(w) = a_{12}(w) = 4\,\sigma^2\,Y$. Here we let:

$$\Pr\left(w + e_1 h + e_2 h \mid w\right) \ = \ \frac{a_{11}(w)/2 + hb_2^{+}(w)}{Q(w)} \ ,$$
$$\Pr\left(w - e_1 h - e_2 h \mid w\right) \ = \ \frac{a_{11}(w)/2 + hb_2^{-}(w)}{Q(w)} \ ,$$
$$\Pr\left(w + e_1 h \mid w\right) \ = \ h\frac{b_1(w) - b_2(w)}{Q(w)} \ ,$$

58

and all other probabilities are zero. In this case the length of the time period is:

$$\Delta t(w) = \frac{h^2}{Q(w)} \quad \text{and} \quad Q(w) = a_{11}(w) + h(|b_2(w)| + b_1(w) - b_2(w))$$

We require

$$a_{11}(w) + hb_1(w) > 0$$

for all probabilities be well defined. Note that for $g = Y$ we have $b_2(w) < b_1(w)$. Also both $b_1$ and $b_2$ are decreasing in $g = Y$ and strictly positive at $g = Y = 0$. Furthermore $a_{11}(w) = 4\sigma^2 Y \geq 0$. Hence, at least for $h$ small enough and $g = Y > 0$ the three probabilities are strictly positive. For $g = Y = 0$ two probabilities are strictly positive.

Let $h = \overline{Y}/(M-1)$ for an integer $M > 1$, and define the grid $\mathbb{W}$ with $M(M+1)/2$ elements as

$$\mathbb{W} = \left\{ (Y_i, g_j) \; : \; Y_i = \frac{i-1}{M-1} \overline{Y} \; , \; i = 1, ..., M \; , \; g_j = \frac{j-1}{M-1} \overline{Y} \; , \; 1 \leq i \leq j \right\}$$

We denote the value function $v : \mathbb{W} \to \mathbb{R}$ as $v_{i,j} = v(Y_i, g_j)$. We define the approximation to the pde in 5 different cases, the interior of $W$ and its boundaries as follows:

(a) For $i = M$ and $j = 1, ..., M$:

$$v_{M,j} = -\Phi/n \; , \tag{A-43}$$

(b) For $i = 3, ..., M - 1$ and $2 \leq j \leq i - 1$:

$$
\begin{aligned}
v_{i,j} = \; & \Delta t_{i,j} \left( \alpha + \left( \frac{n-1}{n} \right)^2 \beta \frac{j-1}{M-1} \right) + \\
& + \; e^{-r\Delta t_{i,j}} \sum_{i'=i\pm1,\,j'=j\pm1} \Pr\left(i', j' \mid i, j\right) v_{i',j'} \\
& + \; e^{-r\Delta t_{i,j}} \sum_{i'=i\pm1,\,j'=j} \Pr\left(i', j' \mid i, j\right) v_{i',j'} \\
& + \; e^{-r\Delta t_{i,j}} \sum_{i'=i,\,j'=j\pm1} \Pr\left(i', j' \mid i, j\right) v_{i',j'} \; ,
\end{aligned}
\tag{A-44}
$$

In this case $\Delta t_{i,j}$ and $\Pr\left(i', j' \mid i, j\right)$ are given by the expressions in case (I).

(c) For $i = 2, ..., M - 1$ and $j = i$:

$$v_{i,j} = \Delta t_{i,j} \left( \alpha + \left( \frac{n-1}{n} \right)^2 \beta \frac{j-1}{M-1} \right) + \quad \text{(A-45)}$$

$$+ \quad e^{-r\Delta t_{i,j}} \sum_{i'=i+1,j'=j+1} \Pr\left(i', j' \mid i, j\right) v_{i',j'}$$

$$+ \quad e^{-r\Delta t_{i,j}} \sum_{i'=i-1,j'=j-1} \Pr\left(i', j' \mid i, j\right) v_{i',j'}$$

$$+ \quad e^{-r\Delta t_{i,j}} \sum_{i'=i+1,j'=j} \Pr\left(i', j' \mid i, j\right) v_{i',j'} \ ,$$

In this case $\Delta t_{i,j}$ and $\Pr\left(i', j' \mid i, j\right)$ are given by the expressions in case (II).

(d) For $i = 2, ..., M - 1$ and $j = 1$:

$$v_{i,1} = \Delta t_{i,1} \, \alpha + e^{-r\Delta t_{i,1}} \sum_{i'=i\pm1,j'=j+1} \Pr\left(i', j' \mid i, 1\right) v_{i',j'} \quad \text{(A-46)}$$

$$+ \quad e^{-r\Delta t_{i,1}} \sum_{i'=i\pm1,j'=j} \Pr\left(i', j' \mid i, 1\right) v_{i',j'}$$

$$+ \quad e^{-r\Delta t_{i,1}} \sum_{i'=i,j'=1+1} \Pr\left(i', j' \mid i, 1\right) v_{i',j'} \ ,$$

In this case $\Delta t_{i,j}$ and $\Pr\left(i', j' \mid i, j\right)$ are given by the expressions in case (I).

(e) For $i = j = 1$:

$$v_{1,1} = \Delta t_{1,1} \, \alpha + e^{-r\Delta t_{1,1}} \sum_{i'=i+1,j'=j+1} \Pr\left(i', j' \mid 1, 1\right) v_{i',j'}$$

$$+ \quad e^{-r\Delta t_{1,1}} \sum_{i'=i+1,j'=j} \Pr\left(i', j' \mid 1, 1\right) v_{i',j'} \ ,$$

In this case $\Delta t_{i,j}$ and $\Pr\left(i', j' \mid i, j\right)$ are given by the expressions in case (II).

$$v(Y, y) = \mathbb{E}\left[ \int_0^{\tau(\bar{Y})} e^{-rt} \left( \alpha - \frac{\beta}{2} y(t) \right) dt - e^{-r\tau(\bar{Y})} \frac{\Phi}{n} \mid Y(0) = Y, \ y(0) = y \right]$$

$$= \mathbb{E}\left[ \int_0^{\tau(\bar{Y})} e^{-rt} \left( \alpha - \frac{\beta(n-1)}{2n} \frac{n}{n-1} y(t) \right) dt - e^{-r\tau(\bar{Y})} \frac{\Phi}{n} \mid Y(0) = Y, \ y(0) = \frac{n-1}{n} \frac{n}{n-1} y \right]$$

$$= \tilde{v}\left( Y, \ y \frac{n}{n-1} \right) \ ,$$

60

where

$$\tilde{v}(Y, g) \;\; \equiv \;\; \mathbb{E}\left[ \int_0^{\tau(\bar{Y})} e^{-rt}\left(\alpha - \frac{\beta\,(n-1)}{2\,n}g(t)\right)dt \; - \; e^{-r\tau(\bar{Y})}\frac{\Phi}{n} \;\middle|\; Y(0) = Y,\; g(0) = g\right]$$

$$\text{for } \; g(t) = \frac{n}{n-1}\; y(t) \text{ with law of motion given by equation (A-42)}$$