

**Income Distribution and the Limits to
Policy Reform
Shock Therapy or Gradualism?**

Vivek H. Dehejia*

Forschungsbericht/
Research Memorandum No. 343
May 1994

* *Address correspondence to:*
Columbia University, Department of Economics
International Affairs Building, Room 1022
420 West 118th Street, New York, NY 10027, U.S.A.
Tel.: + 1 (212) 854-3680
Fax: + 1 (212) 854-8059
E-mail: vd7@columbia.edu

lead to efficiency losses without any compensating gain. More formally, the result follows as a consequence of the first fundamental theorem of welfare economics: if the distorting instruments are removed at the beginning of the problem, and if all other conditions for a competitive equilibrium are satisfied, it follows that the trajectory that the economy follows thereafter will be a Pareto optimum.³ The fact that the equilibrium is dynamic rather than static as in the conventional treatment is purely incidental in this regard.

An important corollary follows from this proposition. Given that shock therapy is the first-best reform policy, it must be that any case for gradualism rests on second-best considerations, i.e., there must exist at least one other distortion in the system that leads to a violation of the assumptions of the first welfare theorem and hence renders shock therapy either infeasible or undesirable. The general wisdom of second-best theory suggests that, once we add additional distortions to the system, we do not in general know what the (constrained) second-best optimum reform policy will look like: it might continue to be shock therapy, it might be gradualism, or it might be something else. Implicitly, therefore, the literature on shock therapy and gradualism considers the implications of adding additional distortions and ascertaining whether shock therapy remains the optimal policy or whether gradualism or some other policy becomes optimal. This literature is briefly reviewed below.

The starting point of this paper is the simple observation that, in the real world, there typically does not exist lump sum income redistribution. This assumption is necessary for the Mussa proposition (and indeed for the first welfare theorem) to hold, since it means that one can legitimately delink the achievement of aggregate economic efficiency from questions of income distribution. To put it another way, it allows a potential Pareto improvement for all agents to be translated into an actual Pareto improvement that leaves no agents worse off and some agents strictly better off than before. In the specific context of the Mussa proposition, it is sufficient for the achievement of social optimality to focus solely on achieving economic efficiency, i.e., to maximize the present discounted value of the economy's final output streams, net of adjustment costs, since lump sum redistribution occurring in the background ensures actual Pareto improvement for all agents. The economic efficiency criterion then implies immediately that shock therapy is the optimal policy.

When one disallows lump sum redistribution, one introduces a distortion which causes the necessary conditions of the Mussa proposition and the first welfare theorem

³ Indeed, in a formal sense, the proof is equivalent to arguing that nonintervention is optimal in an undistorted, competitive economy.

to fail.⁴ The criterion of economic efficiency is no longer innocuous, since there is a legitimate question as to what the appropriate social welfare criterion should be, economic efficiency being only one of several contenders. Selection of a different criterion, especially one sensitive to income distribution, could then possibly lead to a normative case for gradualism.

This, however, is not the approach that I follow in the present paper, since any case for gradualism that an alternative criterion might generate would obviously depend critically on one's acceptance of the criterion, evaluation of which would take us beyond the scope of traditional positive economics.⁵ Rather, accepting the economic efficiency criterion as the criterion for social optimality, I consider whether disallowing lump sum redistribution creates a positive or political economy case for gradualism, for which purpose I need to add a political economy superstructure to the underlying neoclassical economic structure that I use. The answer is a qualified 'yes'.

The approach taken here differs from most of the literature that has evolved since Mussa (1978), which generally introduces major departures from the underlying neoclassical paradigm of Mussa's argument and examines the implications for optimal policy. The interested reader is referred to the comprehensive survey of the existing literature by Rodrik (1993), who carefully categorizes the arguments that reinforce the case for shock therapy and those that generate a case for gradualism.⁶ What is worth noticing is that the papers surveyed typically make assumptions that predispose the analysis either to shock therapy or gradualism. This has two implications. First, it *a fortiori* makes such models inappropriate for analyzing within the context of the given model the choice between shock therapy and gradualism. Second, insofar as these papers take a stand on the shock therapy and gradualism debate, the burden of the argument clearly lies in

⁴ Strictly speaking, one needs to rule out not only explicit lump sum redistribution by a policymaker but also costless and costlessly enforceable bargaining between agents. Else, there would exist a Coasian-type 'bribing' solution in which the winners transfer side payments to the losers, which would replicate in a decentralized fashion the social planner's solution with lump sum redistribution.

⁵ See, however, the concluding section of the paper for a normative reinterpretation of the paper's positive analysis.

⁶ Some interesting papers written on the subject since Rodrik's survey or not included there include Bjorksten (1993a,b), Calvo and Drazen (1993), Caplin and Leahy (1992), Dixit and Rob (1992), Gavin (1993), Martinelli and Tommasi (1994), Nielsen (1994), and Staiger (1994).

respectively.¹¹ Problems (P_X) and (P_Y) can be solved by constructing (current-value) Hamiltonians H_X and H_Y respectively:

$$H_X = F(L_X, K_X) - P_K K_X - q_X I_X + \lambda_X I_X$$

and

$$H_Y = \pi G(L_Y, K_Y) - P_K K_Y - q_Y I_Y + \lambda_Y I_Y$$

where λ_X and λ_Y are the (current-value) costate variables associated with the state variables in problems (P_X) and (P_Y) respectively.

The first-order conditions are as follows for (P_X) :

$$\frac{\partial H_X}{\partial I_X} = -q_X + \lambda_X = 0; \quad (1)$$

$$\frac{\partial H_X}{\partial K_X} = F_K - P_K = 0; \quad (2)$$

$$\frac{\partial H_X}{\partial L_X} = F_L = -\dot{\lambda}_X + r\lambda_X; \quad (3)$$

and for (P_Y) :

$$\frac{\partial H_Y}{\partial I_Y} = -q_Y + \lambda_Y = 0; \quad (4)$$

$$\frac{\partial H_Y}{\partial K_Y} = \pi G_K - P_K = 0; \quad (5)$$

$$\frac{\partial H_Y}{\partial L_Y} = \pi G_L = -\dot{\lambda}_Y + r\lambda_Y. \quad (6)$$

These first-order conditions can be interpreted as follows. Equation (1) equates the marginal cost of hiring an additional work in sector X , q_X , to the marginal benefit, λ_X , which is the shadow value of an additional ‘installed’ worker. Equation (2) equates the marginal product of capital in sector X with its marginal cost in a standard way. Equation (3) is the Euler equation governing the dynamic evolution of λ_X , and it has a standard interpretation as the fundamental equation of asset pricing, which sets the

¹¹ It may seem at first puzzling that only the flow investment rate of labor, rather than its stock, enters the objective function directly. The intuition is that when a firm adds incrementally to its stock of labor, it pays that unit of labor the present discounted value of its lifetime income stream at the point of hiring, analogous to the standard q theory.

normal return ($r\lambda_X$) equal to the sum of the dividend (F_L) and the capital gain ($\dot{\lambda}_X$). Equations (4) to (6) have analogous interpretations for sector Y .

2.2 Retraining Sector

It is time now to elaborate the assumptions made on the mobility of labor, which had been termed a quasi-fixed factor. Suppose that the movement of labor between sectors requires a process of retraining which is undertaken in a retraining sector. Firms in the retraining sector 'buy' labor from sector Y for asset price q_Y and sell it to sector X for asset price q_X (if the direction of flow is from Y to X) and thus earn

$$q \equiv q_X - q_Y$$

per worker trained. What is the cost of retraining a worker? Suppose that retraining employs some of the economy's mobile capital stock as well as a specific factor (say a specialized stock of retraining capital which is not part of the general capital stock) according to a cost of adjustment function

$$C(I_X) \equiv P_K K_I(I_X) \equiv P_K \frac{1}{2} \beta I_X^2,$$

where, following Mussa (1978), I suppose that the capital requirement function and hence the cost of adjustment function is quadratic (in general, any C satisfying $C(0) = C'(0) = 0$, $C'(I_X) > 0 \forall I_X > 0$, and $C''(I_X) > 0 \forall I_X$ will do).¹² Thus, representative retraining firms solve the following optimal control problem, denoted (P_I):

$$(P_I): \max_{\{I_X\}} \int_t^\infty [qI_X - P_K \frac{1}{2} \beta I_X^2] e^{-r(\tau-t)} d\tau.$$

Since the problem has no intertemporal dimension for retraining firms, however, they merely optimize at every instant in time. The first-order condition is the following:

$$q - P_K \beta I_X (\equiv C'(I_X)) = 0, \quad (7)$$

which equates the marginal benefit from moving a worker from sector Y to sector X , q , to the marginal cost, $P_K \beta I_X$.

¹² The assumption of convex adjustment costs, while standard in neoclassical investment theory, has come under attack as of late in the literature; see, for instance, Dixit and Rob (1992). I use it here because it ensures the existence of a well-defined steady state and eliminates the possibility of hysteresis, thus allowing me to present the cleanest argument possible.

2.3 Equilibrium Dynamics

What characterizes equilibrium in this model? Consider first the equilibrium allocation of capital. From equations (2) and (5), and the full employment conditions $L_X + L_Y = \bar{L}$ (which in turn implies that $I_X = -I_Y$) and $K_X + K_Y = \bar{K}$, we have:

$$F_K(L_X, K_X) = \pi G_K(\bar{L} - L_X, \bar{K} - K_X) = P_K,$$

or

$$F_K - \pi G_K = 0, \quad (8)$$

which is the standard static efficiency condition for the allocation of capital.¹³ Indeed, the static efficiency condition holds in this model at every instant in time.

Using equations (1) and (4) and defining

$$\lambda \equiv \lambda_X - \lambda_Y,$$

we get

$$\lambda = q.$$

Subtracting equation (6) from equation (3) and using the equality of λ and q , and defining

$$\Delta \equiv F_L - \pi G_L,$$

we obtain the following equation:

$$\dot{\lambda} = r\lambda - \Delta. \quad (9)$$

This equation is just the fundamental equation of asset pricing for the model, stated in terms of differences of asset values of trained workers in the two sectors. Indeed, we obtain an interesting interpretation of equation (9) by defining a present value shadow price, λ_P , where

$$\lambda_P \equiv \lambda e^{-r(\tau-t)},$$

substituting out λ in terms of λ_P , integrating equation (9) forward, and ruling out explosive bubbles:

$$\lambda_P(t) = \int_t^{\infty} \Delta(\tau) e^{-r(\tau-t)} d\tau,$$

¹³ I have here ignored the use of capital by retraining firms. But, as shall be seen below, in the vicinity of a steady state, the use of capital by retraining firms is negligible, thus making this equilibrium condition locally valid.

which states that the difference in the asset prices of trained workers in sectors X and Y , discounted to the present, is just the present discounted value of the wage differential between sectors, which is what Δ is. For tractability, however, we will continue to work with λ , the current value shadow price, rather than λ_P .

Equation (9) is one of the two fundamental dynamic equations of the model. Rearranging equation (7) and substituting out for q gives the second key equation:

$$\dot{L}_X \equiv I_X = \frac{1}{P_K \beta} \lambda. \quad (10)$$

Together, equations (9) and (10) constitute an autonomous system of two differential equations in two variables, λ and L_X . Steady state conditions are obtained by setting equations (9) and (10) equal to zero and solving for λ and L_X . Setting $I_X = 0$ in equation (10) yields $\lambda^* = 0$ as a steady state condition on λ . Setting $\dot{\lambda} = \lambda = 0$ in equation (9) yields $\Delta = 0$, which implicitly defines a steady-state allocation of labor to sector X (and therefore implicitly to sector Y), L_X^* , such that there exists no wage differential between sectors. Furthermore, it is proved in appendix A.1 that the steady state of the model is locally saddle path stable from an explicit linearization of the model dynamics. This information can be collected in the following proposition:

Proposition 1. *Steady state in the benchmark neoclassical model is characterized by no wage differential between sectors and consequently a split of the labor force between sectors in accord with the HOS model. Furthermore, this steady state is locally saddle path stable. An increase in β , parameter of the cost of adjustment technology, induces a slower approach to steady state along the saddle path.*

Proposition 1 is formally proved in appendix A.1. The steady state and local dynamics for small disturbances around steady state are summarized in Figure 1.

2.4 Efficiency of the equilibrium

For analytical completeness, I now provide a proof of the Mussa proposition, discussed in the introduction, that the decentralized equilibrium to this economy is economically efficient and, under the assumption of lump sum income redistribution, also therefore socially optimal. This can be proved by construction by comparing the solution to the decentralized equilibrium that has been analyzed so far to the central planner's problem. If the two solutions coincide, the proof is complete. This proof, adapted from Mussa (1978), is presented in appendix A.2. It follows *a fortiori* that the speed of adjustment in the decentralized equilibrium is also efficient.

This equivalence of the decentralized and central planner's solutions to the problem in turn proves by construction that the decentralized equilibrium is efficient. This result has a very important interpretation in the context of the discussion of gradualist and shock therapy policy reform in this paper. Suppose that the economy has settled into a steady state characterized by a suboptimally low level of employment in sector X . This could be due to a distortionary tariff, subsidy, or some combination of instruments. Then, suppose at date t the distortions are removed instantaneously, leaving the economy now undistorted but with a level of employment in sector X below its new steady state level. But then this is exactly the situation which has been analyzed, in which the old, distorted steady state is $L_X(t)$ and the new, undistorted steady state is L_X^* . As argued in the introduction, the efficiency of decentralized adjustment thereafter also therefore implies the efficiency of shock therapy - the instantaneous removal of distorting instruments at date t - in the benchmark neoclassical model.¹⁴ The preceding discussion can be summarized in the following proposition:

Proposition 2. *The decentralized competitive adjustment path (and a fortiori the speed of adjustment) is economically efficient in the benchmark neoclassical model. It follows that shock therapy is efficient (and also socially optimal if lump sum redistribution is available) in this model.*

¹⁴ It is essential to be clear on the nature of the thought experiment undertaken. Two experiments which are conceptually distinct but analytically equivalent are considered: (i) suppose an undistorted economy experiences an exogenous terms of trade shift that alters the steady state split of the labor force in accord with the HOS model; and (ii) suppose an economy initially distorted removes instantaneously a set of distorting instruments which has the effect of changing the domestic terms of trade. If in either case the decentralized adjustment path is found to be efficient, then this is what I mean by the efficiency of shock therapy.

3. A Political Economy Model

3.1 The Policy Reform Problem

Suppose at date t the economy in question inherits a relative price of the importable $\pi(t)$ which includes a distorting subsidy (or tariff).¹⁵ Suppose that the objective is to eliminate this subsidy (or tariff), resulting in a new, lower relative price for the importable of π^* .¹⁶

Let us make parametric the comparison between shock therapy and gradualism in the following way. Let us suppose that at time t the policymaker implements an elimination of the distorting instrument and hence generates a price path for π of the following type:

$$\pi(\tau) = \pi^* + (\pi(t) - \pi^*)e^{-\delta(\tau-t)}. \quad (11)$$

In this tractable framework, δ represents the policymaker's control variable. If δ is set to zero, then it is immediate from equation (11) that $\pi(\tau) = \pi(t)$, so that there is no policy reform. As δ goes to infinity, by contrast, $\pi(\tau)$ goes almost instantaneously to π^* , giving the polar case of shock therapy.¹⁷

It should be noted that it is assumed that the policymaker's objective is to ensure the success of the reform, by shock therapy if feasible, and if not by gradualism (more on this below), but failure of the reform is (by hypothesis) not considered acceptable to the policymaker. Intuitively, if reform fails, then in this perfect foresight model, the policymaker cannot attempt reform in the future and hence the economy will be saddled

¹⁵ A question which arises is, how does the government raise the revenue to pay the subsidy or what does it do with the proceeds of the tariff? Since in practice these expenditures and/or revenues are unlikely significantly to affect the changes in income levels in the model, I ignore them. One can think of neutralizing any effect that they might have by supposing that the taxes or rebates arising from the subsidy or tariff, respectively, are borne by the mobile factor, capital, and that these do not affect its opposition to the reform plan.

¹⁶ In particular, if π^* is the exogenous world relative price and ω is the initial distorting subsidy or tariff, then $\pi(t) = \pi^*(1 + \omega)$.

¹⁷ Furthermore, it is assumed that the only instrument available to the policymaker is selection of δ as noted above. I rule out the possibility of other compensation schemes and other types of subsidies and transfer payments, since I shown below that gradualism *suffices* to alleviate the political economy constraint. However, for further discussion of alternatives other than gradualism, see section 4 below.

with the distortion into the infinite future. By contrast, if the distortion is eliminated, even with gradualism, then in the final steady state the economy will be undistorted, a state which the policymaker prefers to the status quo.¹⁸

3.2 A 'Two-Headed' Polity

It is time to spell out the political economy structure of the model. As indicated in the introduction, what is crucial is that workers in sector Y be decisive in influencing the structural adjustment program. This is guaranteed if the unambiguous supporters of reform, i.e., workers in sector X , are not in an absolute majority, since in this case the reform will succeed only with the support of workers in sector Y , recalling that capitalists will unambiguously oppose the reform.¹⁹ For instance, if the three interest groups are of roughly equal mass, this would suffice to make workers in sector Y decisive. Alternatively, suppose for instance that capitalists are of negligible importance in the total mass of voters. Then, workers in sector Y will be decisive if the majority of the labor force is in sector Y at the beginning of the reform program, that is, if the sector that needs to be scaled down has the majority of the economy's total employment at the time that reform is contemplated.

Brief mention should be made of the factor intensity and factor abundance assump-

¹⁸ Of course, this begs the question, why does the distortion exist in the first instance, i.e., why was the policymaker not maximizing economic efficiency then? Or if she was not doing so then, why is she doing so now? This issue is circumvented by simply assuming that the policymaker inherits a distortion, perhaps from a previous regime, at the beginning of the problem.

¹⁹ There is one subtlety here which need not detain us. At the instant of reform, the return to capital rises in terms of the import good but falls in terms of the numeraire export good, whereas in the new equilibrium the return to capital is lower in terms of both goods. Therefore, if capitalists' consumption pattern is biased heavily toward the importable good and there is an extremely slow rate of adjustment, capitalists will perceive a rise in their income in response to a shock therapy reform proposal and will therefore support it. I rule this (unlikely) possibility out and focus exclusively on the case in which capitalists oppose the reform. By contrast for workers, there is no such ambiguity since, regardless of choice of numeraire, real wages go in the same direction. For full generality, then, the model should use a consumption price index as numeraire, rather than the import good as is done here, but doing so adds nothing qualitatively to the analysis while merely complicating the algebra.

tions made, since they affect the political economy. The assumption that sector X is the labor-intensive sector does not entail loss in generality since it merely pins down sector X as the sector guaranteed to benefit from the reform and sector Y the sector we need to worry about. If factor intensities were reversed, we would focus on sector X and be sure that sector Y would support the reform.²⁰ The assumption on factor endowments is important. Were the economy capital-abundant, labor would be the long-run loser of the reform (or terms of trade improvement). Then, workers in the capital-intensive sector would see a rise in their wages in the short run but a fall in the long run, whereas workers in the labor-intensive sector would see lower wages in both the short and long runs. Capitalists, of course, would support the reform, workers in the labor-intensive sector would oppose it, and there would be an ambiguity for workers in the capital-intensive sector. This case could easily be examined. However, the analysis will continue to focus exclusively on the case in which the slowly adjusting factor, labor, is also the long-run beneficiary of the reform, since this in practice seems to me the more interesting case and best seems to capture the possible dilemma that adjustment costs pose for policy reform.

The way that I think of the political economy structure is as follows. Suppose a polity in which the structural adjustment program is devised by an executive (the central planner) who wants to achieve economic efficiency, but who operates under the constraint that her policy has to be ratified by a legislature with three members, one member each for the three groups in the model. The members of the legislature do not vote over all possible reform agendas but have a binary choice over the status quo and the reform agenda proposed by the executive.^{21 22} Then, knowing that she needs to

²⁰ Of course, what we must rule out to keep the analysis tractable is factor intensity reversals in the relevant region.

²¹ This assumption does have bite, since it eliminates the possibility of cycling. In this particular model, however, cycling is not very likely, since for shock therapy to be preferred to gradualism (which would lead to a cycle in which shock therapy is defeated by the status quo, which is defeated by gradualism, which in turn is defeated by shock therapy) we would need that capitalists prefer shock therapy to gradualism (recognizing that workers in the expanding sector prefer shock therapy and those in the contracting sector prefer gradualism). This is unlikely, as capitalists are the steady state losers of the reform; thus, it would be a theoretical *curiosum* if they preferred shock therapy to gradualism.

²² The political economy model used here is closely akin in spirit to the Feenstra

win support of workers in sector Y for the reform to succeed, the executive will solve an optimal control problem in which she maximizes the present discounted value of final outputs, net of adjustment costs, subject to the constraint that workers in sector Y support the reform.²³ This will ensure that the reform agenda receives two out of three votes in a binary choice against the status quo, and the reform will therefore succeed. Furthermore, dynamic inconsistency issues are initially short-circuited (more on this below) by hypothesis by binding both the central planner and the members of the legislature to the equilibrium reform plan selected at date t .²⁴

3.3 An explicit political economy constraint

Following the discussion above, the central planner must now ensure that the reform not diminish the discounted lifetime value of workers in the declining sector, $V_Y(t)$, as compared against the benchmark value enjoyed before the proposed reform.²⁵ But what measures the value of a worker in sector Y ? It will be argued that the correct measure is simply the asset value of an ‘installed’ worker in sector Y , what before had been termed λ_Y (or alternatively q_Y). But what of the possibility of a worker who begins the problem in sector Y ending up in sector X and thus earning λ_X in sector X ? It turns out that the assumption of a competitive retraining sector, made in this paper, implies that the cost of retraining borne by the worker who is switching sectors exactly

and Bhagwati (1982) ‘two-headed’ model of the state, which is not isomorphic with respect to the standard median voter model. Indeed, in this model the assumptions necessary for the median voter theorem to hold are not met and therefore there would not exist a median voter equilibrium to the model. It is crucial for this equilibrium that the legislature has a binary choice determined by the executive. In political scientists’ parlance, it is crucial that there be an ‘agenda-setting’ government for the existence and uniqueness of the political equilibrium studied here.

²³ Of course, there will be similar constraints for the other two groups, but these will be slack in the equilibrium that we are considering and hence can be ignored.

²⁴ A subsequent section of the paper relaxes this assumption.

²⁵ In comparing values as opposed to utilities, I am implicitly ignoring the consumption gain from the relative price change in the case of a tariff reform, a consideration which would further complicate the algebra but not make a qualitative difference. In the case of a production subsidy, consumers always face the world relative price of the importable π^* , so that there is no consumption gain, and the comparison of values is therefore precisely correct.

offsets the increment to lifetime income, leaving the worker no better or worse off than before.²⁶ Formally, suppose that a worker who begins the problem at time t in sector Y switches to sector X at any arbitrary date, say t' . Then, this worker's net discounted lifetime income, $V_Y(t)$, is given by the following expression:

$$V_Y(t) = \int_t^{t'} \pi G_L e^{-r(\tau-t)} d\tau + \int_{t'}^{\infty} F_L e^{-r(\tau-t)} d\tau - C'(t') e^{-r(t'-t)},$$

where $C'(t')$ denotes the marginal retraining cost evaluated and paid by the switching worker to the retraining sector at date t' .

But the second term on the right-hand side just equals $q_X(t') e^{-r(t'-t)}$ by definition, and using equation (7), the second two terms can be replaced by $q_Y(t') e^{-r(t'-t)}$, or, written out in full and factoring through the term $e^{-r(t'-t)}$, we obtain

$$V_Y(t) = \int_t^{t'} \pi G_L e^{-r(\tau-t)} d\tau + \int_{t'}^{\infty} \pi G_L e^{r(\tau-t)} d\tau = \int_t^{\infty} \pi G_L e^{-r(\tau-t)} d\tau,$$

which in turn just equals $q_Y(t)$ or $\lambda_Y(t)$. This may be summarized in the following lemma:

Lemma. *The value of a worker in sector Y is just given by λ_Y , the asset value of an 'installed' worker in sector Y , irrespective of whether the worker moves to sector X at any arbitrary date or remains in sector Y .*

Thus, the political economy constraint recognized by the central planner is given by:

$$\lambda_Y(\tau) \geq \tilde{\lambda}_Y(\tau) \quad \forall \tau \in [t, \infty), \quad (12)$$

where $\tilde{\lambda}_Y(\tau)$ is the value of being in sector Y assuming that the reform is abandoned at date τ (for $\tau = t$, this is just the status quo value of being in sector Y given that reform does not take place). This condition therefore ensures that workers in sector Y never perceive a drop in the value of being in sector Y anywhere along the reform trajectory. Two cases will be considered. In the first case, the case of perfect precommitment, I assume that at date t the policymaker and the legislature can write a binding contract which ties them to the reform agenda as articulated at date t . In the second case, I

²⁶ This result is not surprising in that an asset value in a perfect foresight model such as this one already internalizes all future costs and benefits, including the possibility of incurring an adjustment cost.

relax this assumption by allowing workers in sector Y to defect from this reform path, which they will do at any instant if the value to them of abandoning the reform exceeds the value to them of continuing the reform.²⁷ It is this second case in which issues of time-inconsistency come to the fore.

In the first case, the comparison of values in equation (12) by assumption need be made only at date t . The question arises, when does the political economy constraint fail to bind, that is, when is $\lambda_Y(t)$ at least as big as $\tilde{\lambda}_Y(t)$? Although a fully general answer to the question is not readily available, insight may be gained by examining a linear approximation to equation (12) in the neighborhood of a steady state.²⁸ Performing such an analysis (details of which are contained in appendix A.3) yields the following (local) condition for when the constraint will *not* bind:

$$\frac{(\pi^* G_L^* - \pi(t) G_L(t))}{r} + \frac{\pi^* G_{LL}^* \{L_X(t) - L_X^*\}}{z_2 - r} + \frac{(\pi(t) - \pi^*) G_L^*}{\delta} \geq 0, \quad (13)$$

where z_2 is the stable root of the linearized system (see appendix A.3 for an explicit expression). The intuition in (13) is appealing, in that the three terms represent, respectively, the steady state gain from the reform (which we know is positive by the Stolper-Samuelson theorem and furthermore is constant for reforms of given magnitude), the cost of structural adjustment (which is negative), and the ‘gradualism effect’ (which is positive).

What is the intuition behind this result? Let us look at expression (13) term by term. The first term, as indicated, is simply the annuitized value of the steady state benefit of the reform, which is merely the usual Stolper-Samuelson effect. The second term reflects the presence of adjustment costs. Notice that this term is linear in the discrepancy between the steady state allocation of labor to sector X in the undistorted

²⁷ Throughout the analysis, and in the spirit of the agenda-setting policymaker assumption, it will be assumed that the policymaker has access to a perfect precommitment technology, so that she can always credibly commit to a reform proposal at date t . Departing from this assumption is inconsistent with the existence of a policymaker who can meaningfully set an agenda that will then be voted upon, a central building block of this model.

²⁸ There is an issue as to the propriety of using linear approximations to a dynamic model, which by definition are valid only in the vicinity of a steady state, when in reality reforms and therefore adjustments of finite magnitude are being considered. See the discussion at the end of section 3.3 below.

equilibrium and the lower, distorted allocation of labor to sector X . Notice as well that, as adjustment costs become smaller and therefore less onerous, z_2 becomes more negative and therefore the second term as a whole declines in importance. This is intuitive: the presence of adjustment costs obviously matters less when those costs themselves are less onerous. The final term is what I have dubbed the ‘gradualism effect’, and it essentially reflects the income effect of gradualism on the asset value of workers who begin in the contracting sector. Given a trajectory of reallocation of labor between sectors, gradualism’s effect is to raise the incomes of workers in the declining sector during the transitional period and therefore raise the asset value of being in the declining sector.

It may seem puzzling that the effect of gradualism is unambiguously positive on the value of being in the contracting sector. After all, while gradualism does raise the incomes of workers in the contracting sector, it is a mixed blessing in that it surely also slows down the adjustment process, which implies that the steady state gains of the reform are also pushed further into the future; this latter effect should push down the value of being in the contracting sector. It turns out that, in the vicinity of a steady state (i.e., in the linearized model), the second effect is absent, since the rate of labor reallocation between sectors is a function only of the gap between the initial allocation and the new steady state allocation and of the speed of adjustment in the vicinity of steady state (as is evident from an inspection of equation (A.31) in appendix A.3).

How important is this second effect and can it overturn the positive income effect that gradualism has on the value of being in the declining sector? Numerical simulations²⁹ of the nonlinear dynamics of the model suggest that the second effect is very small in magnitude and therefore that the primary income effect dominates. For instance, in the ‘baseline’ case³⁰ in which $\tilde{\beta} = 3.4$, $\lambda_Y(t) = 49.1358$, which overestimates the true (non-linear) λ_Y by 1.2687. When $\tilde{\beta}$ is set to 10, a very high number in this context, the overestimation increases to 3.87938.

²⁹ These and all other numerical simulations reported in this paper were done on Mathematica 2.0. A copy of the program that generated the simulations is available upon request from the author.

³⁰ In the ‘baseline’ simulations, $\bar{L} = \bar{K} = 100$, F and G are Cobb-Douglas with an exponent on labor of 0.55 and 0.45, respectively, $\pi(t) = 1.005$, $\pi^* = 1$, $r = 0.01$, all of which taken together generate $\beta^* = 3.0009$, and finally $\tilde{\beta} = 3.4$. Of course, there is no loss in generality involved, since all of the parameters (except $\tilde{\beta}$) are normalizations to the extent that they generate β^* . It is then the gap between $\tilde{\beta}$ and β^* which becomes important. The definitions of β^* and $\tilde{\beta}$ are provided in section 3.4 below.

Thus, even when costs of adjustment are extremely high, so that the transitional path of the economy is very long, the error introduced by linearizing the system is relatively small, and the importance of the second effect remains small. It is thus extremely unlikely that gradualism may actually worsen the lot of workers beginning in the contracting sector and therefore overturn the political economy benefit of a gradualist policy. In all of the numerical simulations, it is the primary (income) effect which dominates, as it does in the linearized model of course, and in the sequel, therefore, this is what will be assumed.

3.4 Political economy equilibria in the perfect precommitment case

3.4.1 Optimality of shock therapy

The first question that arises is, when is shock therapy feasible in this setting? Recall that shock therapy coincides with the special case in which δ goes to infinity, which means that the third term in (13) above goes to zero. Then, only two terms are left, the steady-state gain from the reform and the adjustment cost term. Now, it is intuitively apparent (and can be formally established by examining the explicit expression for z_2 in appendix A.3) that the speed of adjustment becomes less rapid as the cost of adjustment (represented parametrically by β) increases; furthermore, given that the everything else is held constant, z_2 is monotonic in β . Thus, as adjustment costs become more onerous (higher β), the speed of adjustment becomes slower (z_2 becomes more negative) and therefore the cost of adjustment term becomes more negative.³¹ In fact, it is evident that there exists a critical speed of adjustment z_2^* (and an underlying critical adjustment cost parameter β^*) which solves (13) with equality when the third term is set to zero. At this speed of adjustment, workers in sector Y are just indifferent between supporting and opposing the shock therapy reform. If adjustment costs are just slightly more onerous, hence adjustment just slightly slower, the sum of the two terms will become strictly negative and workers in sector Y will oppose the reform. Conversely, for slightly less onerous adjustment costs and hence slightly more rapid adjustment, the sum of the two terms will become strictly positive and workers in sector Y will support the reform. This is shown in Figure 2, in which the expression in (13) is compactly notated $\Phi(\beta)$.

³¹ Of course, we are not doing an experiment in which we are considering altering β in the same economy, which would be illegitimate, but rather we are comparing economies each of which has a different β .

Algebraically, z_2^* is given by

$$z_2^* = r - r \left(\frac{\pi^* G_{LL}^* \{L_X(t) - L_X^*\}}{\pi^* G_L^* - \pi(t) G_L(t)} \right) \quad (14)$$

and in turn β^* is given by

$$\beta^* = \frac{\Delta_L^*}{-z_2^*(z_2^* - r)P_K^*}, \quad (15)$$

where I have used the definition of z_2 as given in appendix A.3

The implications of this analysis for the feasibility of shock therapy reform are immediate. Shock therapy reform will be feasible in an economy whose adjustment costs are less onerous than β^* and will be infeasible in an economy whose adjustment costs are more onerous than β^* . In the former case, since shock therapy is feasible, and it is also economically efficient (as was proved in section 2.4 above), it therefore will be implemented by the agenda-setting policymaker. Intuitively, if the period of adjustment is sufficiently short, the steady state gains of the reform will outweigh the costs of adjustment to affected workers, and shock therapy will be feasible and therefore 'politically optimal'.³² We can collect this information in the following proposition:

Proposition 3. *There exists a critical adjustment cost parameter β^* such that if the economy's actual β , $\tilde{\beta}$, is greater than this, shock therapy reform will hurt workers (i.e., have a net effect on lifetime income which is negative) in sector Y and hence be infeasible. By contrast, if $\tilde{\beta}$ is less than this β^* , sector Y workers will enjoy a net benefit from the reform and hence shock therapy will be feasible and therefore 'politically optimal', i.e., shock therapy will be selected as the optimal reform plan by the agenda-setting policymaker and will be ratified by the legislature.*

3.4.2 Optimality of gradualism

If indeed the economy is characterized by small adjustment costs, shock therapy will be feasible and optimal and hence gradualism will not be proposed by the policymaker. However, if shock therapy is not feasible (because β exceeds β^*), will gradualism serve to make the policy reform feasible? The answer is 'yes'. Suppose that the economy's β is some $\tilde{\beta}$ which exceeds β^* . Then, we know from section 3.4.1 above that the first two terms in (13) will be negative. It follows therefore that there exists a finite and

³² Of course, in this case the issue of time-inconsistency does not arise since by construction the reform is accomplished at one fell swoop at date t .

positive value of δ , say δ^* , which ensures that the third term exactly offsets the first two terms and satisfies (13) with equality. Figure 3 depicts δ^* as a function of $\tilde{\beta}$. If δ is set even slightly higher than δ^* , then (13) will be strictly negative and workers in sector Y will oppose the reform, and similarly if δ is slightly lower than δ^* , (13) will be strictly positive and workers in sector Y will support the reform.

Algebraically, δ^* is given by

$$\delta^*(\tilde{\beta}) = \frac{(\pi(t) - \pi^*)G_L^*}{-\Psi(\tilde{\beta})}, \quad (16)$$

where $\Psi(\tilde{\beta}) \leq 0$ is defined to equal the first two terms in equation (13) and where by construction $\Psi(\beta^*) = 0$.

Knowing this, the policymaker can therefore announce a reform program *ex ante* characterized by equation (11) and with a value of δ just epsilon below δ^* . Workers in sector Y will compute a slight increase in their discounted lifetime incomes and will therefore support the reform in coalition with workers in sector X and against the capitalists. The reform will therefore succeed. This therefore will be the equilibrium gradualist reform policy selected by the policymaker and approved by the legislature.³³

From Figure 3 and equation (16), notice that δ^* is drawn as decreasing in $\tilde{\beta}$, which is perfectly intuitive (and can easily be verified). With more onerous adjustment costs, a more gradual reform program will be necessary to ensure support of workers in sector Y and hence ensure feasibility of the reform. Furthermore, notice that δ^* asymptotes vertically at β^* , since whenever the economy's β , $\tilde{\beta}$, is equal to β^* or less, shock therapy is feasible and therefore δ^* should be set to infinity; and it asymptotes along the horizontal axis, since for a bigger and bigger $\tilde{\beta}$, a lower and lower δ^* is required to ensure feasibility of the reform.

The key result of this section can be stated in the following proposition:

Proposition 4. *If adjustment costs are sufficiently onerous and shock therapy is not feasible, then there exists a gradualist reform program, defined parametrically by δ^* , which ensures that workers in sector Y will support the reform and hence that the reform will be feasible. This then defines the agenda-setting policymaker's equilibrium reform program which will be ratified by the legislature.*

³³ Of course, the policymaker will not select a δ that is smaller than δ^* by a finite amount since doing so prolongs the adjustment process and needlessly imposes an additional distortion on the economy. However, see section 3.5 below for a qualification of this in the case without precommitment.

3.5 Political economy equilibria in the absence of precommitment

In the absence of precommitment, the constraint in equation (12) must hold for every date τ , not just at the initial date t . What does this do the case for gradualism? As it turns out, it serves only to strengthen it. To see this, consider the following analysis of this case.

What we would like is an analogue to expression (13) which holds at every date τ . It is given by the following:

$$\begin{aligned} & \frac{(\pi^* G_L^* - \pi(\tau) G_L^*(\tau))}{r} + \frac{\pi^* G_{LL}^*}{z_2 - r} \{L_X(\tau) - L_X^*\} \\ & + \frac{(\pi(\tau) - \pi^*) G_L^*}{\delta} - \frac{\pi(\tau) G_{LL}^*(\tau) \{L_X(\tau) - L_X^*(\tau)\}}{z_2 - r} \geq 0. \end{aligned} \quad (17)$$

Evidently, the expression in (17) is very similar to that in (13). The only difference (aside from the fact that we now evaluate the quantities at every τ and not just at t) is the addition an extra term, the fourth term. This fourth term, which looks very similar to the adjustment cost term, captures the fact that, if workers decide to abandon the reform at some date τ , they do not immediately reap the steady state income level associated with the relative price prevailing at that date, $\pi(\tau)$, but rather must adjust to that level just as if they were adjusting to any other terms of trade shock. The effect of this fourth term, *ceteris paribus*, is to lower the appeal of abandonment, since abandonment in effect entails a shock therapy-type adjustment vis-a-vis a new, distorted and still suboptimal allocation of resources $L_X^*(\tau)$.

Given that δ has been selected equal to δ^* in accord with the rationale in section 3.4, in the case with perfect precommitment, will the time-consistency constraint, expression (17), ever bind? Denoting the expression in (17) as $\Phi(\tau)$, we do know that, for δ^* , $\Phi(0) = 0$, by construction, and also $\Phi(\infty) = 0$, since the reform is already complete and therefore abandonment is no longer a possibility. In general, however, it is not possible to say what happens for all τ in between, since as time elapses each of the four terms is changing and in different directions. The first term becomes less positive, since the steady state benefit from continued reform is now lower. This, *ceteris paribus*, favors abandonment. The second term becomes less negative, however, since the adjustment costs of continued reform are also lower, which, *ceteris paribus*, favors continuation of the reform. The third term becomes less positive since, as time elapses, the benefit of gradualism declines (since more of the reform has been carried out), which, *ceteris paribus*, favors abandonment. The fourth term in itself is difficult to sign and even more difficult to interpret.

Of course, if the ‘good’ terms (i.e., those which favor reform over abandonment) in (17), which are the second and possibly the fourth term, dominate the ‘bad’ ones, which are the first, third, and possibly also the fourth term, then δ^* , selected without paying heed to the time-consistency issue, will remain politically feasible even under the possibility of abandonment, and all of the analysis of section 3.4 will carry over without modification. However, it is perfectly conceivable that the ‘bad’ terms will dominate and lead to abandonment of the reform at the originally selected δ^* .

The response to the challenge of the time-consistency problem that I take is two-pronged. First, while it is difficult if not impossible to come up with an analytical argument as to when the originally selected δ^* will work and when it will not, it is possible to obtain numerical simulations of (17). Such numerical simulations indicate that, except in cases in which the gap between $\tilde{\beta}$ and β^* is very small, so that δ^* is accordingly very large, the condition is never violated, so that the originally selected δ^* remains politically feasible throughout the reform trajectory and so that the fear of abandonment never in fact materializes. For instance, in the ‘baseline’ simulations, in which $\beta^* = 3.0009$, time-consistency always holds so long as $\tilde{\beta}$ is no smaller than 3.2635. Figure 4 depicts the ‘baseline’ simulation, with $\tilde{\beta} = 3.4$ as usual, in which $\Phi(\tau)$ shows an inverted U-shape which furthermore is skewed toward the left, suggesting that the terms which work toward continuation of the reform dominate early on but that gradually those favoring abandonment begin to take over; importantly, however, the curve never dips below the horizontal axis, so that abandonment never becomes optimal.

The second response to the challenge of time-consistency is the following analytical argument. Since all of the terms in expression (17) are finite and bounded, it follows that, even in those cases in which condition (17) fails for δ^* , it is always possible to make condition (17) hold by sufficiently decreasing δ , since the only term that explicitly contains δ is inversely related to δ . Thus, as δ is pushed lower and lower, the third term becomes arbitrarily large and in the limit approaches infinity. Thus, for all feasible trajectories, a very large third term, induced by a small enough δ , will always dominate. Denote by δ^{**} the largest δ which meets the time-consistency constraint.³⁴ By construction, $\delta^{**} < \delta^*$ (else, we would not have gone through the exercise of finding it).

Thus, the analysis of the case without precommitment falls into subcases. In the event that the originally selected δ^* ensures that condition (17) always holds, time-

³⁴ Such a δ^{**} could in specific cases be constructed by the simple algorithm of starting with δ equal to δ^* and successively lowering it by a small, fixed increment until condition (17) holds for every τ .

consistency holds and therefore the analysis of section 3.4 above carries through as before. However, if this is not the case, time-consistency binds, and it has been argued that there exists a $\delta^{**} < \delta^*$ which ensures that the time-consistency condition will be satisfied. It follows that in those cases in which the time-consistency condition binds, satisfying the constraint in fact requires the policymaker to be even more gradual (i.e., select a smaller δ) than in the case with perfect precommitment.

We can summarize the information presented above in the following proposition:

Proposition 5. *In the case without perfect precommitment, issues of time-inconsistency strengthen the case for gradualism rather than weaken it. If the time-consistency constraint is not binding, then Proposition 4 always holds. If the constraint does bind, however, then the policymaker will have to select a δ , $\delta^{**} < \delta^*$, which implies that the reform will have to be even more gradual than in the perfect precommitment case.*

3.6 Are smaller reforms politically easier?

An issue which is closely allied to the time-consistency question and which has remained in the background should be mentioned briefly at this juncture. This concerns the question as to whether smaller reforms are politically more feasible than larger reforms. In fact, this formulation is a bit loose. The more rigorous way of asking the question is: how is the critical level of adjustment costs at which shock therapy is just feasible affected as we alter the magnitude of reform? If the sustainable level of adjustment costs falls for larger reforms, then it is less likely that any given economy will have low enough adjustment costs to sustain a shock therapy reform. Thus, while for any given economy shock therapy reform is either feasible or not, depending on the gap between $\tilde{\beta}$ (the economy's actual β) and β^* , if we suppose that there exists a continuum of potential economies, one for each $\tilde{\beta}$, then fewer of these economies will find reform feasible if β^* decreases.

Interestingly, this question is indirectly linked to the question we examined in section 3.4 above, namely, for a given δ , will the time-consistency condition (17) hold? If we eliminate the fourth term from expression (17), the remaining terms capture the effects on the value of remaining in sector Y as against abandonment at each τ , where now implicitly we are imagining that the initial π , is not $\pi(t)$ but rather some lower $\pi(\tau)$, recalling that π declines monotonically along its trajectory. Obviously, there is the same ambiguity discussed above in section 3.5, since the smaller reform offers both a smaller steady state gain and accordingly smaller benefit from gradualism (the first and third terms) but also has smaller costs of adjustment (the second term).

Suggestive as the above consideration is, for a frontal assault on the problem we must check how β^* (and accordingly δ^*) change as $\pi(t)$, the initially distorted price of good Y , changes. Here, as in the time-consistency issue explored above, an analytical solution is hard to come by, and accordingly I once again resort to numerical simulation to get a flavor of what the answer might be. The ‘baseline’ simulation is presented in Figure 5, which should be interpreted as follows. In the simulation presented, π^* is normalized to unity, so that as we move rightward along the horizontal axis (the origin being unity along the horizontal axis) we are considering larger and larger initial distortions, $(\pi(t) - \pi^*)$. The vertical axis depicts β^* , which you will recall is that level of the adjustment cost parameter at which shock therapy is just feasible. Intuitively, the curve asymptotes at infinity as $\pi(t)$ approaches unity, since at $\pi(t) = 1 (= \pi^*)$ there is no reform to do and hence any level of adjustment costs are politically sustainable. The curve slopes down monotonically, which indicates that larger reforms are politically more difficult – the sustainable level of the adjustment cost parameter becomes smaller and smaller. For a very large initial distortion, the curve appears to asymptote at zero, suggesting that only arbitrarily small levels of adjustment costs are sustainable, implying that no actual economy will find such reform politically sustainable under shock therapy. Furthermore, notice that the curve appears to fall away from the vertical axis very rapidly, i.e., even for $\pi(t)$ slightly greater than 1, β^* is a relatively small positive number.

What is the intuition behind the result? Evidently, when comparing a larger reform as against a smaller one, the fact that the adjustment costs are larger for a larger reform dominates the fact that the larger reform offers a larger steady state gain (and accordingly larger scope for gradualism), and accordingly the larger reform is politically more difficult to carry through under shock therapy. While this ranking of the magnitudes of the effects need not always go this way – in principle it could go either way and I have not found a way analytically to rank order these effects *a priori* – all of the numerical simulations look like the ‘baseline’ simulation in which the adjustment cost effect dominates the other effects, making smaller reforms politically more feasible. Accordingly, I offer the following as a conjecture rather than a proposition.

Conjecture. *In numerical simulations of the model presented in this paper, the critical level of the adjustment cost parameter at which shock therapy reform is just feasible, β^* , is smaller for larger reforms. This implies that it is less likely that any given economy will find shock therapy reform feasible for a larger initial distortion, since it will be less likely that the economy’s actual adjustment cost parameter, $\tilde{\beta}$, will be smaller than β^* ,*

which is required for the feasibility of shock therapy (from Proposition 3).

There is one nice implication of this conjecture worth brief mention.³⁵ If we imagine that the initial distortion is inherited from the ‘mistakes’ of past regimes, it follows then that the initial distortion represents the ‘accumulated’ mistakes of the past. In this case, the lesson for policymakers is that, if a small distortion has crept in, it is politically far easier to tackle it right away rather than to let the distortion accumulate, as it were, which makes the eventual reform politically much more difficult.

4. Alternatives to Gradualism

The core of this paper’s analysis, as presented in section 3 above, has considered gradualism as the only alternative policy to shock therapy. Accordingly, whenever shock therapy reform is politically infeasible, it follows that the only alternative policy available to the agenda-setting policymaker is gradualism. Indeed, one of the chief arguments of this paper has been that there will always exist a gradualist alternative when shock therapy reform proves to be infeasible.

But what of other alternative policy measures which may be tried if shock therapy is politically infeasible? In this section of the paper, I briefly consider two alternatives, pre-announced (optimally delayed) shock therapy and subsidizing worker retraining. As shall be seen, I conclude that gradualism, for different reasons in each case, remains the most appealing alternative.

4.1 Pre-announced (optimally delayed) shock therapy

This alternative to gradualism involves the policymaker selecting an optimal date $T \geq t$ in the future, at which date shock therapy will take place. T is to be selected to be the earliest date which will ensure that the announcement of shock therapy at that date in the future will be politically feasible (i.e., will leave workers in the declining sector just indifferent in value terms, as in the analysis in section 3 above).

While it is analytically onerous formally to construct this alternative, the intuition behind it is perfectly clear. By announcing at date t that shock therapy will take place at date T in the (possibly quite distant) future, in this perfect foresight model and barring issues of time-inconsistency, it is conceivable that the value of being in the declining sector will not fall, since this optimally delayed shock therapy offers the prospect of a

³⁵ I thank Andreas Wörgötter for suggesting this to me.

steady state gain in the future and a long transitional path in which to prepare for the move to the other sector. Indeed, some retraining and moving between sectors would presumably begin at date t as the announcement occurs.

There are two issues that arise in considering the alternative. The first is the simple question as to how this policy and gradualism are rank ordered in terms of their efficiency effects. Both policies involve efficiency losses, and in true second best fashion it is impossible to come up with a unique rank ordering by (qualitative) analytical argumentation. Furthermore, without an explicit analytical expression for the trajectory under pre-announced shock therapy, it is obviously impossible to compare the two policies parametrically.

The more serious issue is that the policy of pre-announced shock therapy is prone to very serious time-inconsistency problems which are potentially far more troublesome than those which might bedevil the gradualism alternative. For consider the following alternative scenario.³⁶ Suppose that, at date t , when the announcement is made, no one believes that at date T anything will happen, so no one starts moving between sectors. Then, when date T finally arrives, no one will have moved, and therefore by construction shock therapy will not be feasible, accordingly the reform will not be carried out, and the expectation that nothing will happen will have been validated. This alternative trajectory in which no one believes the policymaker and nothing happens is also consistent with agents' rationality and perfect foresight, as is the more optimistic scenario sketched above. Evidently, we are dealing in this case with an issue of multiple equilibria, which is certainly not very appealing insofar as pre-announced shock therapy is considered a sensible alternative to gradualism.³⁷ While the preceding discussion has admittedly been somewhat impressionistic, it suggests that gradualism is a more appealing alternative than pre-announced shock therapy, which seems prone to very serious time-inconsistency issues.

4.2 Subsidizing worker retraining

A more sensible alternative and one which is likely to have more backers than pre-announced shock therapy is the policy of subsidizing worker retraining. The rationale behind this alternative is easy enough to see. Since the key problem in this economy

³⁶ I thank Shang-Jin Wei for suggesting this to me.

³⁷ Recall that it was possible to ensure by an appropriate selection of δ that time-inconsistency would be avoided in the gradualism case. No such comparable argument can be made here.

is apparently the presence of costly adjustment, why not subsidize the retraining of workers, thereby speeding up the adjustment process and accordingly ensure that shock therapy reform will be politically feasible?

In the context of the model employed in this paper, it is easy to show that the smallest subsidy to the retraining sector which makes shock therapy reform just feasible is given by the following expression,

$$\sigma^* = 1 - \frac{\beta^*}{\tilde{\beta}}, \quad (18)$$

where σ^* is the subsidy. The intuition is transparent: implementing a subsidy of σ^* ensures that the economy's $\tilde{\beta}$ net of the subsidy is just equal to β^* , the adjustment cost parameter which makes shock therapy just feasible, since

$$\beta^* = \tilde{\beta}(1 - \sigma^*)$$

by construction. Thus, it is politically feasible for the agenda-setting policymaker to propose shock therapy reform at date t , even though $\tilde{\beta} > \beta^*$, if only she also proposes an optimally selected retraining subsidy σ^* as part of the package.

This alternative to gradualism is clearly a far more serious contender than pre-announced shock therapy and it has the appeal that it allows shock therapy reform to be carried out. It is also interesting in that it represents a political economy rationale for subsidizing worker retraining, whereas the more commonly articulated arguments for retraining which one comes across usually rely on the assumption of credit market imperfections which prevent workers from borrowing against future income to finance retraining today.³⁸

However, it must be remembered that, like gradualism, subsidizing worker retraining brings with it an efficiency loss. Under the gradualist alternative, the efficiency loss stems from the fact that the distorting subsidy or tariff remains in place after date t , which leads to the standard neoclassical production loss from a misallocation of the economy's resources during the transitional period (plus a consumption loss, only in the case of a tariff). Indeed, it is precisely because of the production loss that the reform is undertaken in the first instance, although obviously gradualism entails a smaller loss than the status quo, since under gradualism the distortion is eventually eliminated.

³⁸ Of course, it is possible that these usual reasons are the rhetoric and the political economy argument the true rationale for subsidizing retraining!

The merit of the subsidizing retraining alternative is that it makes shock therapy feasible, and hence it avoids the production losses associated with gradualism. Instead of this, however, it substitutes another loss which stems from the fact that there is excessively rapid redeployment of the economy resource's as compared to the efficient speed of adjustment. This is precisely why the policy works, of course: it allows the decentralized economy to mimic the adjustment path of an economy for which shock therapy is just feasible, but doing so puts an additional burden on the economy since too many workers retrain too quickly from the efficiency point of view, so that adjustment costs are suboptimally high.

There is an interesting contrast between gradualism and subsidizing worker retraining as regards the channels through which they operate. The former boosts the incomes of workers in the declining sector and accordingly (as a second order effect) slows the rate at which workers retrain, whereas the latter speeds the adjustment process as its primary effect and thus ensures that workers in the declining sector reap the steady state benefits of the reform soon enough to be at least indifferent to reform.

Again, we are in a second best world in which there is no *a priori* presumption that one or the other policy will entail a larger or a smaller efficiency loss. At first blush, one's intuition might favor the retraining alternative, since it seems to tackle the source of the distortion, excessively slow adjustment to a terms of trade change, head on, whereas gradualism might seem a roundabout way to solve the problem. Subsidizing worker retraining might therefore seem to be consonant with the Bhagwati-Ramaswami-Johnson 'targeting' principle. In fact, this intuition is misguided, since the real source of the problem is not excessively slow adjustment but, as argued, the inability of the winners from reform to compensate the losers and hence the political infeasibility of carrying out shock therapy reform coupled with lump sum redistribution. Once this is understood, it becomes evident that any departure from shock therapy plants us firmly in a second best world in which (almost) anything goes.

Once again, as it is impossible analytically to rank order the policies, I resort to numerical simulations to compare the two policies under specific assumptions on the model's parameters, according to the criterion of which one entails a smaller discounted efficiency loss. Specifically, the criteria used are as follows:

$$\Theta_G = \int_t^{\infty} \{(F_G + \pi^* G_G) - (F_W + \pi^* G_W)\} e^{-r(\tau-t)} d\tau$$

and

$$\Theta_R = \int_t^{\infty} \left\{ \frac{\sigma^* \lambda^2}{2P_K \tilde{\beta} (1 - \sigma^*)} \right\} e^{-r(\tau-t)} d\tau.$$

Θ_G represents the present discounted value of the momentary production loss induced by gradualism (where the subscript G denotes the values of F and G under the gradualist regime and W denotes that F and G are being evaluated at world prices π^*), and Θ_R is the present discounted value of the momentary losses due to excessively rapid adjustment under subsidized retraining, where it easily shown that the momentary loss due to subsidizing retraining takes the form within the curly brackets (this comes from substituting out for I_X in $C(I_X)$ from equation (10)).

The results of the numerical simulations are striking, since they reveal that, in all the cases tried, the efficiency loss from gradualism is smaller than the efficiency loss from subsidizing worker retraining. For instance, in the 'baseline' case (with $\tilde{\beta} = 3.4$), $\Theta_G = 0.359522$ and $\Theta_R = 0.408482$, implying a differential of 0.0489605 in favor of gradualism. Intuitively, the efficiency losses from gradualism are small because the neo-classical production loss from a small subsidy or tariff is small, and when appropriately integrated and discounted these costs remain small. By contrast, the efficiency losses from subsidizing worker retraining are relatively large, and grow larger the larger is $\tilde{\beta}$. Some sense of the relative magnitude of the efficiency losses induced by the two policies may be garnered from Figure 6, which plots (for the 'baseline' case) the difference between the momentary efficiency loss of the retraining subsidy less the momentary efficiency loss of gradualism, appropriately discounted, as against alternative values for $\tilde{\beta}$ ranging in this simulation from 3.4 to 10 and as against time, running in this simulation from 0 to 150 (after which we are virtually at steady state for all practical purposes). As the figure reveals, the retraining subsidy does especially badly compared to gradualism at the beginning of the reform trajectory, after which retraining starts to do a lot better; as expected, the net difference in efficiency losses asymptotes to zero as time becomes large. For each $\tilde{\beta}$, it turns out that integrating the discounted difference in efficiency losses yields a positive number (as noted above for the 'baseline' value of $\tilde{\beta}$), indicating that subsidizing retraining yields a bigger net loss in terms of efficiency than gradualism.

Thus, while subsidizing worker retraining is not prone to the obvious difficulties that beset pre-announced shock therapy, it seems that in numerical simulation gradualism does better in terms of the standard efficiency criterion of which delivers the smaller discounted loss to national income.³⁹

³⁹ The one important caveat is that I consider only the production and not the consumption losses in the case of gradualism, so that strictly speaking the numerical simulations indicate that gradualism does better than subsidizing retraining in the case of a production subsidy. In the case of a tariff, the additional consumption losses might tip

5. Conclusion

In this concluding section, I will not summarize the findings of the paper, but rather turn to the normative interpretation of the model (as promised in the introduction). As noted before, the economic efficiency criterion is unobjectionable in those contexts in which it is assumed that the policymaker has available to her lump sum redistributive instruments, since in such a case the objectives of achieving efficiency and achieving a fair distribution (however society defines this) of the efficient output are delinked. However, in the absence of lump sum redistribution, the criterion ceases to be innocuous, although it has the merit of (perhaps) being the least unpalatable amongst the many alternatives. Indeed, the use of the criterion in this model can be defended on grounds of parsimony, since the analysis in the paper shows that even a policymaker who does not intrinsically care about distribution but who faces a legislature with a veto interest in the reform proposal is forced to take distributional considerations into account.

However, one can imagine situations in which a policymaker who does not have access to lump sum redistribution may choose an alternative social welfare criterion, even if unimpeded by political constraints. Some of these possible alternative criteria may also generate a case for gradualism. I have found one alternative criterion which makes a nice benchmark alternative to the efficiency criterion since it replicates the results of the positive analysis. This alternative I dub a 'quasi-Rawlsian' social welfare criterion, since it supposes that the policymaker has the objective of achieving economic efficiency, subject to the condition that the 'worst off' members of society not be made any worse off due to the reform. If we (perhaps gratuitously) equate the 'worst off' with workers in the declining sector, then this quasi-Rawlsian policymaker, even if unimpeded by a legislature, will select exactly the same gradualist reform program in exactly the same situations as an efficiency-driven policymaker whose choices are politically constrained.

To conclude, let me step beyond the strict confines of the model to consider what lessons the analysis presented here might have for policymakers in concrete situations. The spirit of the model suggests that, whenever politically influential groups stand to be hurt a lot in the short run from shock therapy reform proposals, even if they are the long run gainers from reform, it might be wise for policymakers to implement gradualist reform instead.⁴⁰ Else, they might find that these influential groups will (rationally from their own point of view) block reform altogether, leaving the policymaker stymied and

the balance in favor of subsidizing retraining.

⁴⁰ Indeed, in the real world, capital market imperfections, deviations from rational expectations, or finite planning horizons will make individuals place a greater weight on

the economy stuck in the quagmire of the status quo.

short run losses than on long run gains than they would under perfect capital markets, perfect foresight, and an infinite planning horizon, which are the assumptions in this model, making individuals even more sensitive than the model suggests to income losses early on in a proposed reform.

APPENDIX

A.1 Linearized dynamics of the neoclassical model

Linearizing equations (9) and (10) around a steady-state, we obtain the following system,

$$\begin{bmatrix} \dot{\lambda} \\ \dot{L}_X \end{bmatrix} = \begin{bmatrix} r & -\Delta_L^*(\cdot) \\ \frac{1}{\beta P_K^*(\cdot)} & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ L_X - L_X^* \end{bmatrix}, \quad (\text{A.1})$$

where it is understood that the partial derivatives are evaluated at the steady-state. The characteristic roots of the system are given by:

$$r_1 = \frac{r + \sqrt{r^2 - 4\Delta_L(\cdot)\frac{1}{\beta P_K^*(\cdot)}}}{2}, r_2 = \frac{r - \sqrt{r^2 - 4\Delta_L(\cdot)\frac{1}{\beta P_K^*(\cdot)}}}{2}. \quad (\text{A.2})$$

Recalling that $\Delta_L(\cdot)$ is negative and $\frac{1}{\beta P_K^*(\cdot)}$ is positive, we see that r_2 is negative, while r_1 is positive, thus establishing the saddle-path stability property that I appealed to in the text and confirming that r_2 is the stable root of the system. The eigenvector associated with the stable eigenvalue is

$$\begin{bmatrix} 1 \\ e_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{1}{r_2 \beta P_K^*(\cdot)} \end{bmatrix} \quad (\text{A.3})$$

and since the second element of the eigenvector is negative, we know that the saddle-path is downward-sloping in the (λ, L_X) plane. Inspection of the equation for r_2 confirms that an increase in β reduces the absolute value of r_2 and thus induces a slower approach to the steady state, thus establishing Proposition 1.

Explicit solutions to the linearized system are given by the following equations:

$$\lambda(\tau) = \lambda(t)e^{r_2(\tau-t)}, \quad (\text{A.4})$$

$$L_X(\tau) = L_X^* + [L_X(t) - L_X^*]e^{r_2(\tau-t)}, \quad (\text{A.5})$$

where

$$\lambda(t) = \frac{-\Delta_L^*\{L_X(t) - L_X^*\}}{r_2 - r} \quad (\text{A.6})$$

is the initial condition for λ (defining, in effect, the equation of the saddle path) and is determined by the requirement that the constant term multiplying the unstable root in the equilibrium trajectories for λ and L_X must equal zero.

A.2 Efficiency of the equilibrium

Consider an economy in which the fictitious central planner maximizes the economy's discounted output stream subject to the technological and physical constraints that the economy faces. Call this problem (P):

$$(P) : \max_{\{I_X, K_X, K_Y\}} \int_t^\infty [F(L_X, K_X) + \pi G(L_Y, K_Y)] e^{-r(\tau-t)} d\tau$$

$$s.t. L_X + L_Y = \bar{L};$$

$$K_X + K_Y + K_I \leq \bar{K};$$

$$K_I = \frac{1}{2} \beta I_X^2;$$

$$\dot{L}_X = I_X;$$

$$L_X \geq 0;$$

$$L_Y \geq 0;$$

$$L_X(t), L_Y(t) = \bar{L} - L_X(t), \pi, r \text{ given};$$

where the notation is as before. Again, (P) is solved by constructing the (current-value) Hamiltonian H :

$$H = F(L_X, K_X) + \pi G(\bar{L} - L_X, K_Y) - \gamma(K_X + K_Y + \frac{1}{2} \beta I_X^2 - \bar{K}) + \Lambda I_X.$$

The necessary conditions for a maximum are given by:

$$\frac{\partial H}{\partial I_X} = -\gamma \beta I_X + \Lambda = 0; \quad (A.7)$$

$$\frac{\partial H}{\partial K_X} = F_K - \gamma = 0; \quad (A.8)$$

$$\frac{\partial H}{\partial K_Y} = \pi G_K - \gamma = 0; \quad (A.9)$$

$$\frac{\partial H}{\partial L_X} = F_L - \pi G_L = -\dot{\Lambda} + r\Lambda; \quad (A.10)$$

$$\dot{L}_X \leq 0 \text{ if } L_X = \bar{L}; \quad (A.11)$$

$$\dot{L}_X \geq 0 \text{ if } L_X = 0. \quad (A.12)$$

Here, Λ is the (current-value) costate variable associated with the state variable L_X . Equations (A.8) and (A.9) are the standard static efficiency conditions for the

allocation of capital which must hold at every point in time. But it is evident that these combined using the full employment condition on capital are identical to the decentralized efficiency conditions for the allocation of capital given by equation (8), except that what we called P_K before, the equilibrium domestic cost of capital, is replaced by γ , the Lagrange multiplier on the capital stock constraint, which can nicely be interpreted as the shadow value of relaxing the constraint on the capital stock incrementally. In a distortionless competitive market, this is equal to the equilibrium cost of capital. Equation (A.7) equates the marginal cost from reallocation of labor to its (shadow) marginal benefit. Equation (A.10) is the Euler equation governing the evolution of this (shadow) marginal benefit.

Equations (A.11) and (A.12) are boundary conditions. Equations (A.7) and (A.10), slightly rewritten, give the system of two differential equations which constitute the equilibrium dynamics of the central planner's problem:

$$I_X = \frac{1}{\gamma\beta}\Lambda; \quad (\text{A.13})$$

$$\dot{\Lambda} = r\Lambda - \Delta; \quad (\text{A.14})$$

where once again the compact notation $\Delta \equiv F_L - \pi G_L$ has been used. Comparing the system of equations (A.13) and (A.14) with the system of equations (10) and (9), it is evident that they are identical; what has been called λ in the decentralized problem corresponds to Λ in the central planner's problem.⁴¹ Since $L_X(t)$, the initial split of the labor force, is given by history and is identical in the two problems, the two solutions coincide exactly so long as the initial values $\lambda(t)$ and $\Lambda(t)$ coincide, since thereafter the differential equations governing their behavior are identical. But recalling Proposition 1, the existence of a saddle path guarantees that, given the initial $L_X(t)$, the initial $\lambda(t)$ must be chosen to jump onto the saddle path to ensure stable convergence to the steady state. Specifically, as shown above in appendix A.1, the saddle path is defined by the following equation:

$$\lambda(t) = \frac{-\Delta_L^* \{L_X(t) - L_X^*\}}{r_2 - r}, \quad (\text{A.15})$$

where r_2 , of course, is the stable root to the system, and equation (A.5) has been used to substitute into equation (A.4). But then this completes the proof, since two variables governed by the same differential equation and with the same initial condition behave identically.

A.3 Linearized dynamics of the political economy model

⁴¹ In the background, an identical split of the capital stock between sectors along the decentralized and optimal paths is guaranteed by the equality of P_K and γ .

The system of equations now is very similar that analyzed in appendix A.1 above. The only difference is that I wish to keep λ_X and λ_Y separate variables (important for the political economy discussions in section 3 of the text) and there is an explicit path for $\pi(\tau)$. The system is given by:

$$\dot{\lambda}_X = r\lambda_X - F_L(L_X), \quad (\text{A.16})$$

$$\dot{\lambda}_Y = r\lambda_Y - \pi G_L(\bar{L} - L_X), \quad (\text{A.17})$$

$$\dot{L}_X = \kappa(\lambda_X - \lambda_Y), \quad (\text{A.18})$$

and

$$\pi(\tau) = \pi^* + (\pi(t) - \pi^*)e^{-\delta(\tau-t)}, \quad (\text{A.19})$$

where the notation is as before and furthermore we define for brevity $\kappa \equiv (1/P_K^*\beta)$.

Denote the steady state equilibrium values as λ_X^* , λ_Y^* , and L_X^* , in which $\lambda_X^* = \lambda_Y^* \equiv \lambda^*$ and L_X^* depends only upon π^* .

Linearizing around the steady state yields:

$$\dot{\lambda}_X = r(\lambda_X - \lambda^*) - F_{LL}^*(L_X - L_X^*), \quad (\text{A.20})$$

$$\dot{\lambda}_Y = r(\lambda_Y - \lambda^*) - (\pi - \pi^*)G_L^* + \pi^*G_{LL}^*(L_X - L_X^*), \quad (\text{A.21})$$

and

$$\dot{L}_X = \kappa(\lambda_X - \lambda^*) - \kappa(\lambda_Y - \lambda^*). \quad (\text{A.22})$$

Using equation (A.19), we write equations (A.20) - (A.22) in matrix form as follows:

$$\begin{bmatrix} \dot{\lambda}_X \\ \dot{\lambda}_Y \\ \dot{L}_X \end{bmatrix} = \begin{bmatrix} r & 0 & -F_{LL}^* \\ 0 & r & \pi^*G_{LL}^* \\ \kappa & -\kappa & 0 \end{bmatrix} \begin{bmatrix} \lambda_X - \lambda^* \\ \lambda_Y - \lambda^* \\ L_X - L_X^* \end{bmatrix} + \begin{bmatrix} 0 \\ -(\pi(t) - \pi^*)G_L^* e^{-\delta(\tau-t)} \\ 0 \end{bmatrix}. \quad (\text{A.23})$$

It is known that stable solutions are of the form:

$$\lambda_X - \lambda^* = c_1 e^{z_2(\tau-t)}, \quad (\text{A.24})$$

$$\lambda_Y - \lambda^* = c_2 e^{z_2(\tau-t)} + \frac{(\pi(t) - \pi^*)}{\delta} G_L^* e^{-\delta(\tau-t)}, \quad (\text{A.25})$$

and

$$L_X - L_X^* = c_3 e^{z_2(\tau-t)}, \quad (\text{A.26})$$

where z_2 is the stable root (negative eigenvalue) of the 3x3 matrix of partial derivatives in equations (A.23) and (c_1, c_2, c_3) is the corresponding eigenvector.

To obtain the eigenvalues, we solve the following equation:

$$\begin{vmatrix} z - r & 0 & F_{LL}^* \\ 0 & z - r & -\pi^* G_{LL}^* \\ -\kappa & \kappa & z \end{vmatrix} = 0. \quad (\text{A.27})$$

We obtain:

$$z(z - r)^2 + (z - r)\kappa F_{LL}^* + \pi^* G_{LL}^* \kappa(z - r) = 0.$$

The root $z_1 = r > 0$ factors out, leaving the quadratic equation:

$$z^2 - rz + \kappa(F_{LL}^* + \pi^* G_{LL}^*) = 0,$$

where

$$\Delta_L^* = F_{LL}^* + \pi^* G_{LL}^* < 0,$$

recalling our definition of Δ :

$$\Delta \equiv F_L - \pi G_L.$$

The two other roots are given by:

$$z_2 = \frac{r - \sqrt{r^2 - 4\kappa\Delta_L^*}}{2}, \quad z_3 = \frac{r + \sqrt{r^2 - 4\kappa\Delta_L^*}}{2},$$

where evidently $z_3 > 0$ and hence $z_2 < 0$ is the stable root that we are seeking.

Furthermore, the constants (c_1, c_2, c_3) must satisfy:

$$\begin{bmatrix} z_2 - r & 0 & F_{LL}^* \\ 0 & z_2 - r & -\pi^* G_{LL}^* \\ -\kappa & \kappa & z_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = 0. \quad (\text{A.28})$$

But it is easily seen that

$$c_1 = \frac{-F_{LL}^*}{z_2 - r} c_3, \quad c_2 = \frac{G_{LL}^*}{z_2 - r} c_3,$$

whereas c_3 must obey:

$$c_3 = L_X(t) - L_X^*.$$

Our final solutions are as follows:

$$\lambda_X(\tau) = \lambda^* - \frac{F_{LL}^*}{z_2 - r} (L_X(t) - L_X^*) e^{z_2(\tau-t)}, \quad (\text{A.29})$$

$$\lambda_Y(\tau) = \lambda^* + \frac{\pi^* G_{LL}^*}{z_2 - r} (L_X(t) - L_X^*) e^{z_2(\tau-t)} + \frac{(\pi(t) - \pi^*) G_L^*}{\delta} e^{-\delta(\tau-t)}, \quad (\text{A.30})$$

and

$$L_X(\tau) = L_X^* + (L_X(t) - L_X^*) e^{z_2(\tau-t)}. \quad (\text{A.31})$$

But then using the fact that $\lambda^* = \pi^* G_L^* / r (= F_L^* / r)$, setting $\tau = t$ in equation (A.30) (to obtain the initial value on the saddle path) and substituting into equation (12) yields equation (13) in the text.

REFERENCES

- Bhagwati, Jagdish (1994a). Shock Treatments, *The New Republic*, March 28, 39 - 43.
- Bhagwati, Jagdish (1994b). Reply to Sachs, *The New Republic*, April 25, 5.
- Bjorksten, Neil (1993a). Trade and Structural Unemployment, mimeo (Prague, Central European University).
- Bjorksten, Neil (1993b). Public Policy Alternatives to Economies Suffering Trade Generated Structural Unemployment, mimeo (Prague, Central European University).
- Calvo, Guillermo A. and Allan Drazen (1993). Uncertain Duration of Reform: Dynamic Implications, mimeo (Washington, DC, International Monetary Fund).
- Caplin, Andrew and John Leahy (1992). The Danger of Gradualism in Economic Policy, mimeo (New York, NY, Columbia University).
- Desai, Padma (1993). Ease Up on Russia, *The New York Times*, December 10, A35.
- Desai, Padma (1994). Opening the door, *The New Republic*, February 28, 4 - 5.
- Dixit, Avinash and Rafael Rob (1992). Switching Costs and Sectoral Adjustments in General Equilibrium with Uninsured Risk, mimeo (Princeton, NJ, Princeton University), forthcoming in *Journal of Economic Theory*.
- Feenstra, Robert C. and Jagdish Bhagwati (1982). Tariff Seeking and the Efficient Tariff, in: J. Bhagwati, ed., *Import Competition and Response* (Chicago, IL, University of Chicago Press), 245 - 258.
- Gavin, Michael (1993). Unemployment and the Economics of Gradualist Policy Reform, mimeo (New York, NY, Columbia University).
- Gligorov, Vladimir (1994). Gradual Shock Therapy, East European Series No. 1 (Vienna, Institute for Advanced Studies).
- Martinelli, Cesar and Mariano Tommasi (1993). Sequencing of Economic Reforms in the Presence of Political Constraints, mimeo (Los Angeles, CA, University of California, Los Angeles).

- Mussa, Michael (1978). Dynamic Adjustment in the Heckscher-Ohlin-Samuleson Model, *Journal of Political Economy* 86, 775 - 791.
- Mussa, Michael (1982). Government Policy and the Adjustment Process, in: J. Bhagwati, ed., *Import Competition and Response* (Chicago, IL, University of Chicago Press), 73 - 120.
- Mussa, Michael (1986). The Adjustment Process and the Timing of Trade Liberalization, in: A. Choksi and D. Papageorgiou, eds., *Economic Reform in Developing Countries* (Oxford, Basil Blackwell), 68 - 124.
- Nielsen, Claus Kastberg (1994). Gradualist Trade Reform: Virtue or Vice?, Department of Economics Memo 1994-13 (Aarhus, University of Aarhus).
- Rodrik, Dani (1993). The Positive Economics of Policy Reform, *American Economic Review* 83, Papers and Proceedings, 356 - 361.
- Sachs, Jeffrey (1993). Poland's Jump to the Market Economy (Cambridge, MA, MIT Press).
- Sachs, Jeffrey (1994a). Betrayal, *The New Republic*, January 31, 14 - 18.
- Sachs, Jeffrey (1994b). Not-so-dismal science, *The New Republic*, April 25, 4 - 5.
- Staiger, Robert W. (1994). A Theory of Gradual Trade Liberalization, NBER Working Paper Series No. 4620 (Cambridge, MA, National Bureau of Economic Research).

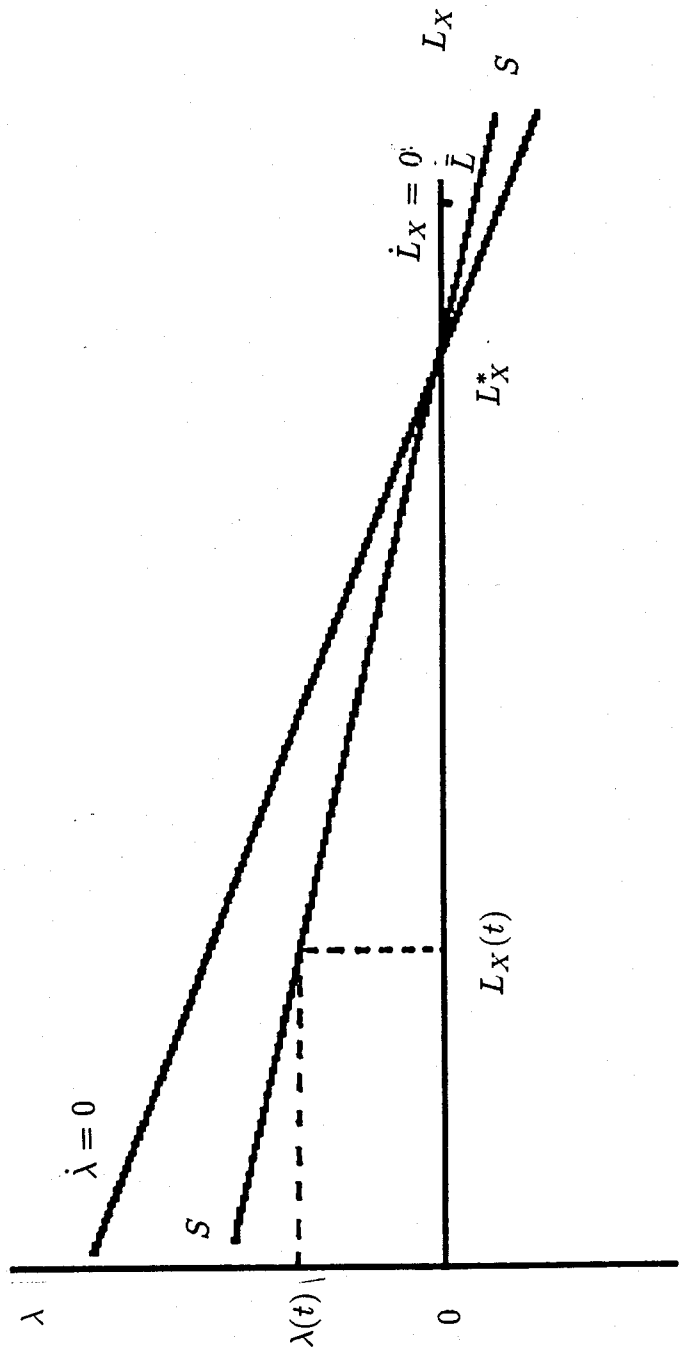


Figure 1. Equilibrium dynamics in the neoclassical model.

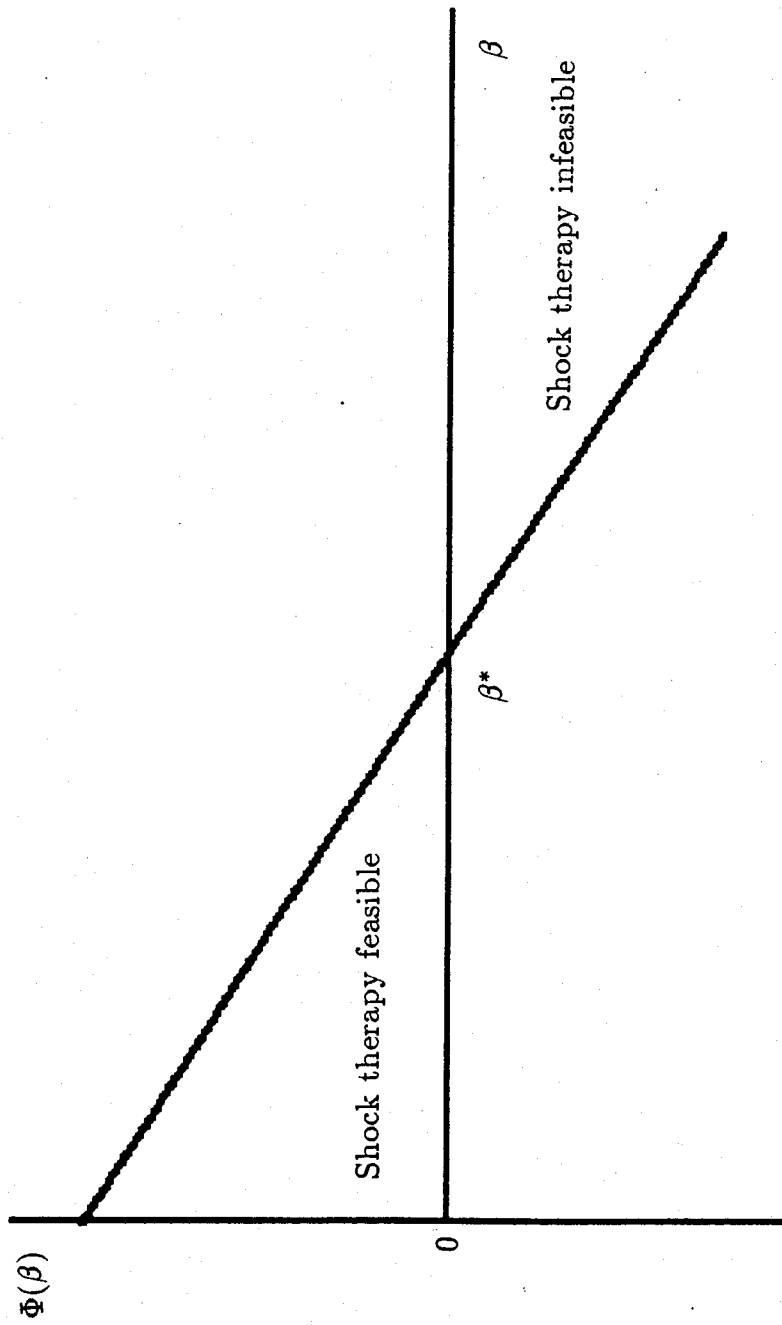


Figure 2. Determination of the critical cost of adjustment parameter (β^*) for the feasibility of shock therapy.

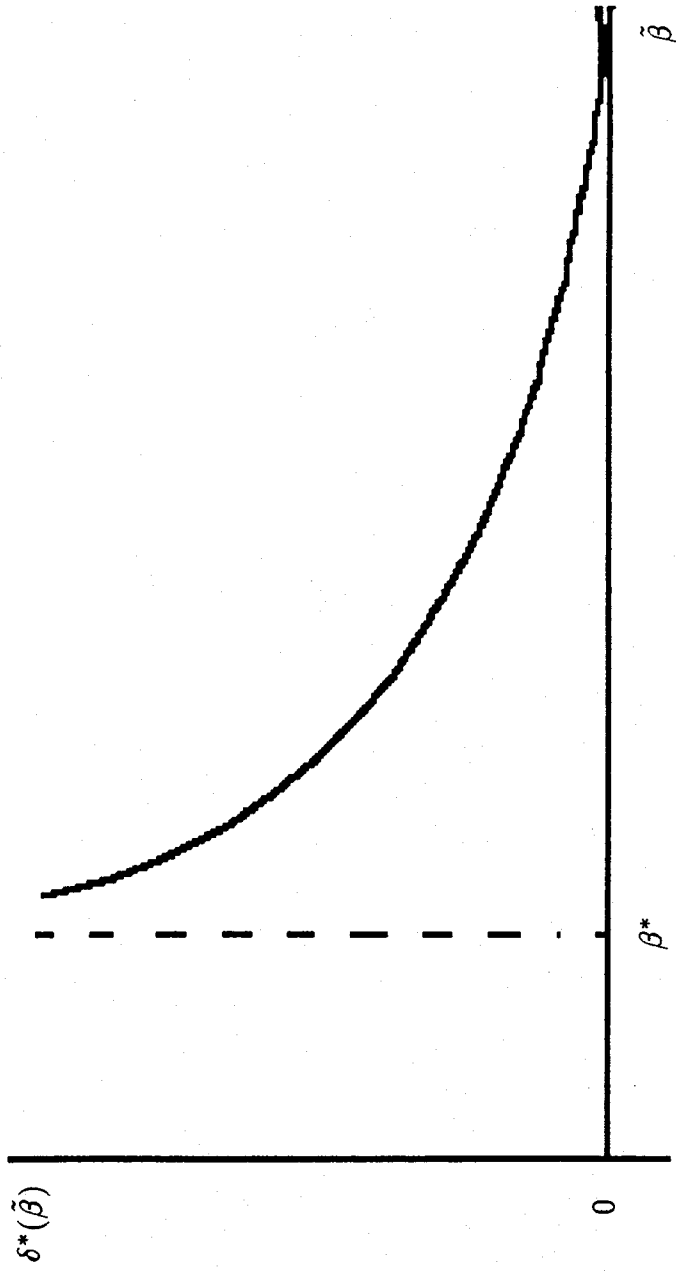


Figure 3. Determination of the optimal gradualist reform agenda (δ^*) as a function of the cost of adjustment parameter(β).

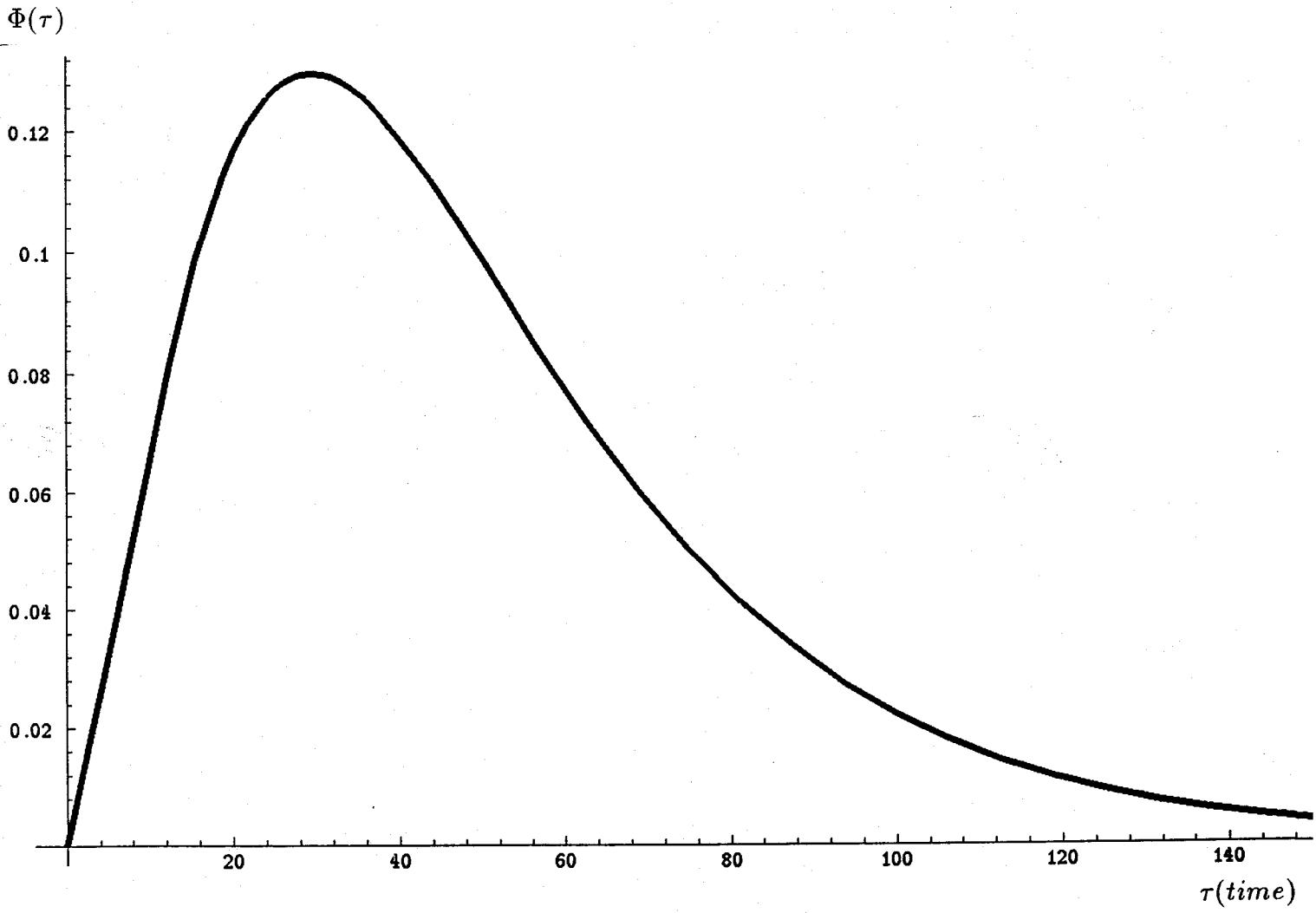


Figure 4. Value to workers in the declining sector of continuing reform minus value of abandonment in numerical simulation.

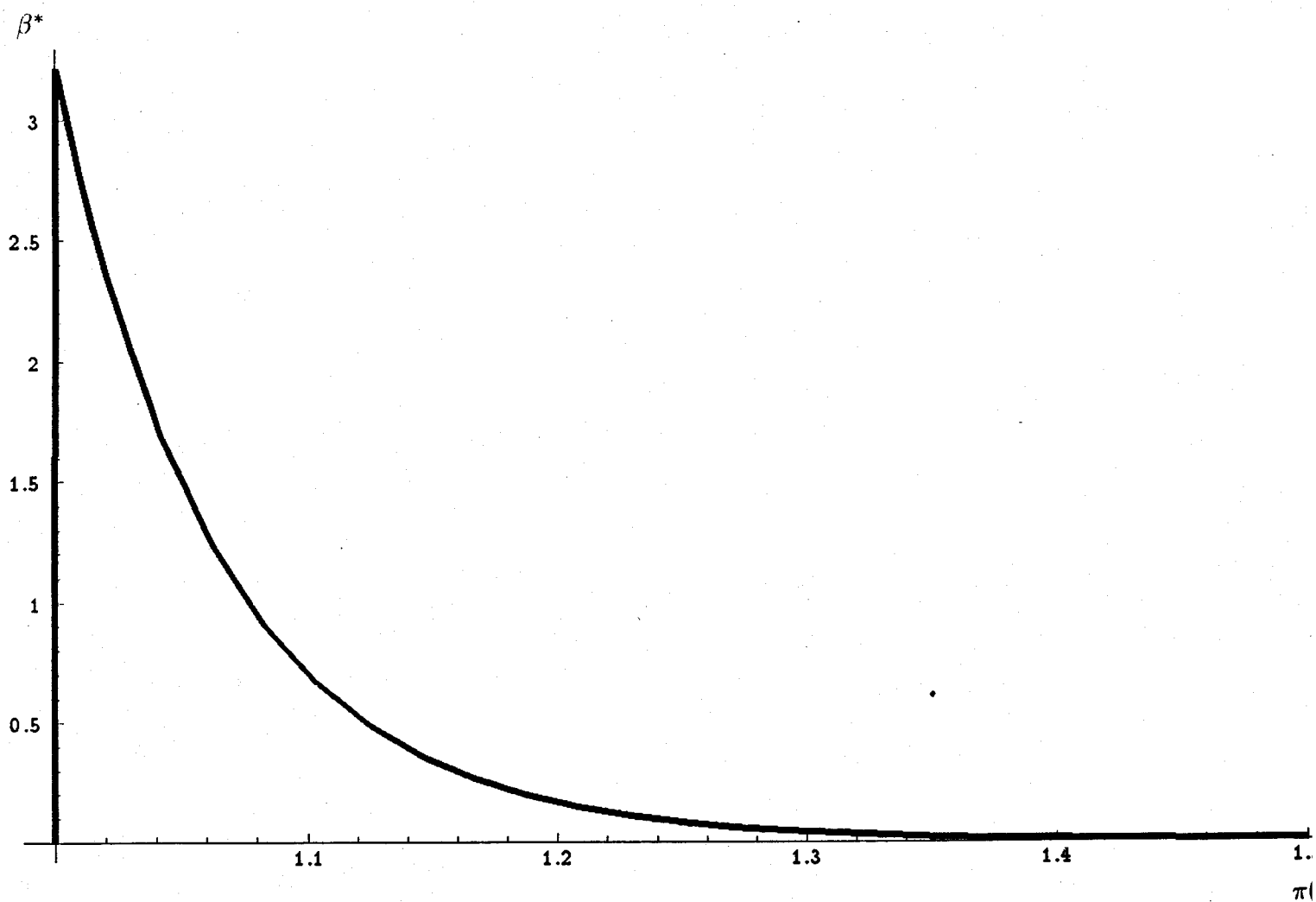


Figure 5. The critical cost of adjustment parameter (β^*) as a function of the magnitude of reform in numerical simulation.

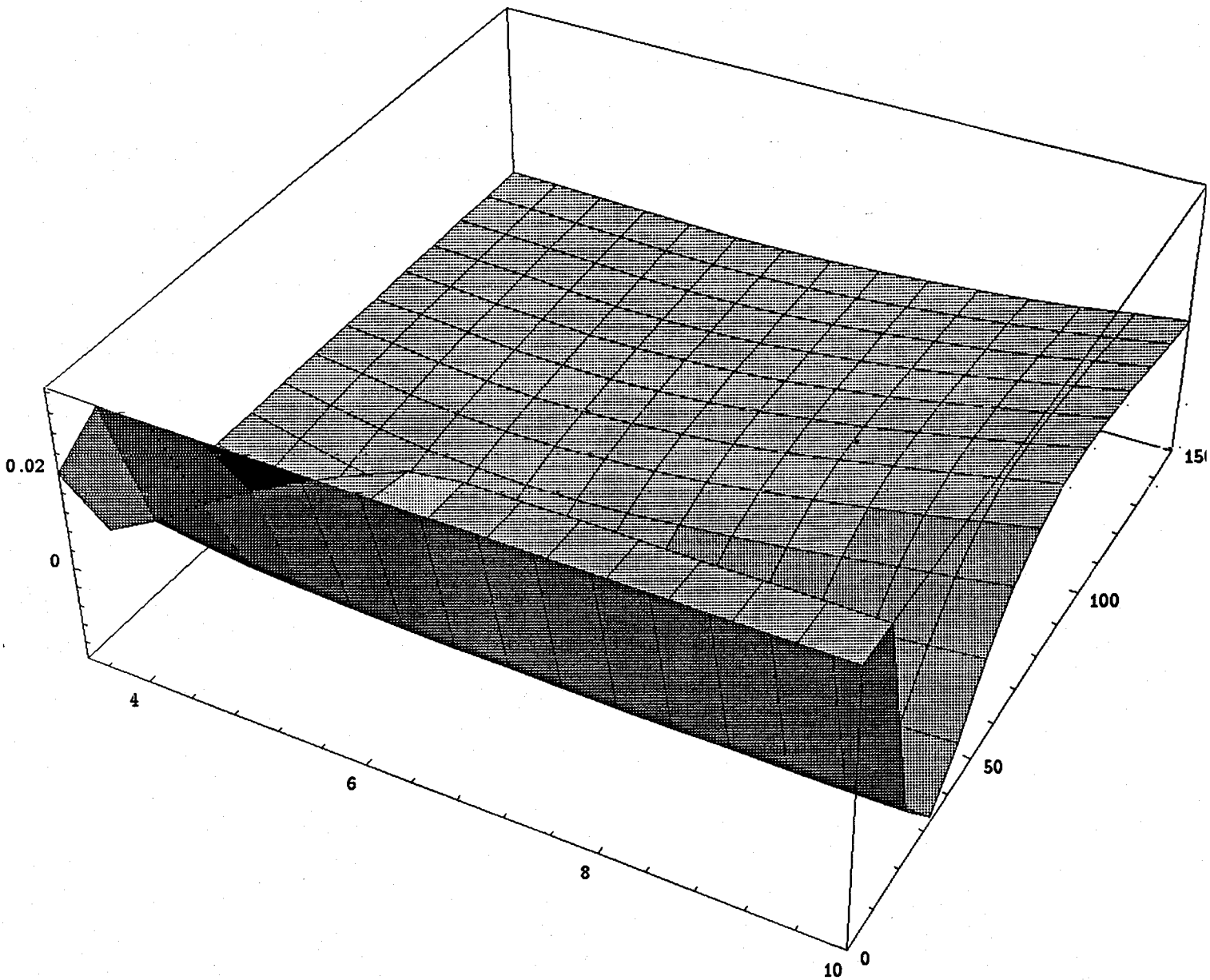


Figure 6. The discounted momentary difference between the efficiency cost of a retraining subsidy and the efficiency cost of gradualism.