TRANSACTIONAL MICROSTRUCTURE AND MACROECONOMIC PERFORMANCE†

The "Fundamental Transformation" in Macroeconomics

By RICARDO J. CABALLERO AND MOHAMAD L. HAMMOUR*

When factors enter into joint production, they often develop a degree of "specificity" with respect to each other. Their value in the context where they have come to produce exceeds their outside value. Committing a production factor in such a relationship is a form of specific investment, whose sunk component may have technological as well as institutional origins.

By itself, this "irreversibility" has important and extensively studied implications for investment decisions. But, when combined with *contracting* problems, it acquires an entirely new dimension. When two factors enter into a joint-production relationship, specificity creates *ex post* quasi-rents that need to be protected through *ex ante* contracting. Unless complete and enforceable contracts are available, those quasi-rents may be *appropriated* by the other factor. This transformation from an *ex ante* competitive setting to an *ex post* bilateral monopoly—what Oliver Williamson terms the "fundamental transformation"—is a central building block in the modern economic theory of institutions (e.g., Benjamin Klein et al., 1978; Williamson, 1979).

The problem of appropriability spreads unprotected quasi-rents throughout the economy and, in a partial-equilibrium setting, leads to underinvestment. In general equilibrium, the market system will adjust to help partly compensate the appropriated factors, providing a highly inefficient macroeconomic "solution" to the unresolved microeconomic contracting problem. Appropriability affects major aspects of the macroeconomy. It results in misallocation, underutilization, and unemployment of the economy's productive factors; it hampers growth by depressing the incentives to replace what is outdated and to utilize fully the economy's resources; it disrupts macroeconomic adjustment, by inducing a wedge between timid creation and excessive destruction of the old system; and it exacerbates downturns by "elastifying" the cyclical response of inelastic factors.

I. Microeconomics

The prototypical macroeconomic example of specificity we choose to focus on is that of the relationship between capital and labor. Other examples are also important and affect such diverse relations as that between an entrepreneur and his external sources of finance (e.g., Oliver Hart and John Moore, 1994) or that between an upstream firm and its downstream customers (as discussed extensively in the microeconomic literature on vertical integration).

In the presence of specificity, the central unit of analysis is the joint-production relation between the factors. We embody this relation in the concept of a "production unit" which, we assume, combines in fixed proportions one unit of capital and one unit of labor. Because of specificity, production units cannot be costlessly reshuffled. It is therefore meaningful to analyze both the *creation* decision, by which factors choose to enter into a production relation, and the *destruction* decision, by which factors within an existing production unit choose whether to continue producing jointly or not.

† *Discussants:* Steven J. Davis, University of Chicago; Mark Gertler, New York University; Abhijit Banerjee, Massachusetts Institute of Technology.

* Caballero: Department of Economics, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139, and NBER; Hammour: Capital Guidance, 32 Avenue Hoche, 75008 Paris, France. We thank Olivier Blanchard and Steven Davis for helpful comments.

We outline a minimalist model to help structure our discussion. For a more complete treatment, we refer the reader to Caballero and Hammour (1996a, b, c). The units of analysis that arise from the notion of specificity provide the natural setting for a Schumpeterian view of a macroeconomy with ongoing creation and destruction (see e.g., Caballero and Hammour, 1994). In this short paper, however, we reduce our treatment of dynamics to a minimum. We model destruction and creation decisions during a single period. We assume that factors can either be part of a production unit in the "joint-production" sector, or remain in their "autarky" sectors. The period starts with preexisting production units as well as a supply of uncommitted units of capital and labor. First, the factors in all preexiting units decide whether to produce jointly in the coming period or to separate and join the uncommitted factors; second, all uncommitted factors either form new production units or remain in autarky; finally, production of the economy's unique consumption good takes place in the joint-production and autarky sectors.

Creation.—We denote by y the production revenue of a newly created production unit, and by w_k and w_ℓ the reward of capital and labor in autarky (all in terms of the consumption good). In partial equilibrium, the *efficient* decision is to create a production unit as long as

$$y \geq w_k + w_\ell . \tag{1}$$

This rule would, in fact, be followed as long as factors of production can freely and costlessly move in and out of production units.

Specificity.—Such freedom of movement rarely holds in reality. Let us take the extreme irreversible-investment view, where once capital is locked into a specific relationship, it is lost if the relationship breaks up. Although it can also be due to such institutional factors as organized labor or legislated firing costs, we view specificity here as a purely technological phenomenon. As a result, while the *ex ante* opportunity cost of capital is w_k, its *ex post* opportunity cost is zero. On the other extreme, we assume that labor develops no specificity with respect to the relationship, so w_ℓ represents both its *ex ante* and its *ex post* opportunity cost.

The Fundamental Transformation.—If capital and labor were able to precommit to their *ex post* terms of trade, they would follow the efficient entry rule. But when complete contracts are costly to write and enforce, agents' ability to precommit is limited, and a Williamsonian "fundamental transformation" arises: the *ex ante* competitive relationship between capital and labor turns into an *ex post* bilateral monopoly. Assuming that no precommitment is possible, the specific quasi-rent s from a production unit is the difference between the unit's revenue and its factors' *ex post* opportunity costs:

$$s \equiv y - w_\ell . \tag{2}$$

Following the Nash bargaining solution for sharing the unit's revenue, we assume that each factor gets its *ex post* opportunity cost plus half of the unit's bargaining-surplus s. Using the superscript j to denote payments to factors in new joint-production units, we have

$$w_k^j = \frac{s}{2} \tag{3}$$

$$w_\ell^j = w_\ell + \frac{s}{2} . \tag{4}$$

Underinvestment.—It is clear from equations (3) and (4) that labor will willingly participate in production as long as the surplus from joint production is nonnegative. Capital, the factor that is vulnerable to appropriation, requires more. Its share of the *ex post* surplus must be enough to justify forfeiting its *ex ante* opportunity cost:

$$\frac{s}{2} \geq w_k .$$

Replacing (2) into this expression yields the entry rule more explicitly:

$$y \geq 2w_k + w_\ell \tag{5}$$

which states that a unit's revenue must not only cover its factors' opportunity costs, as in the efficient condition (1), but must also generate enough rents for capital to recover its *ex ante* opportunity cost after being appropriated. This increased-revenue requirement generally gives rise to an *underinvestment* problem, a result most clearly stated by Paul Grout (1984) in this context.

Destruction.—The other decision of interest to us is that of a preexisting production unit with production revenue x that must decide whether to remain in operation or shut down and free its nonspecific resources. The privately efficient rule is to exit if the opportunity cost of the two factors of production cannot be covered. Since capital is fully specific, its opportunity cost is zero. On the other hand, the opportunity cost of labor in a preexisting unit is $pw_\ell^j + (1 - p)w_\ell$, where p is the probability of finding employment in joint production (see below). The exit rule is therefore

$$x < pw_\ell^j + (1 - p)w_\ell . \tag{6}$$

In the context of bilateral bargaining, the question of the private efficiency of separations arises. When there is a "nontransferability" problem, due, for example, to asymmetric information (as in Robert Hall and Edward Lazear [1984], for example) or to a minimum-wage constraint, the exit decision may be privately inefficient (see e.g., Roger Myerson and Mark Satterthwaite, 1983). We abstract away from this possibility and assume for simplicity that exit decisions are privately efficient. As we will show, this is in sharp contrast with their social inefficiency at the macroeconomic level.

II. Macroeconomics

In our simple setup, macroeconomic analysis consists of understanding the equilibrium determination of factor rewards in autarky and joint production, and the corresponding utilization of the economy's factor endowments. For this, we assume that the initial aggregate supply of uncommited factors is 1 for both capital and labor. To simplify the determination of equilibrium quantities, we also assume that preexisting production units are of infinitesimal mass. We assume that, in its autarky sector, capital can be turned one-to-one into the consumption good, so $w_k = 1$. On the other hand, labor experiences decreasing returns in its autarky sector, resulting in a joint-production labor supply function:

$$w_\ell = 2L^{1/\eta} \qquad \eta > 0 \tag{7}$$

where L measures joint-production employment. The supply of capital into joint production is therefore perfectly elastic, while the supply of labor is partly inelastic. Finally, we restrict our analysis to values of $y \in (2, 3)$, where both the efficient and inefficient economies exhibit positive but not full employment in joint production.

A. *The Efficient Economy*

If precommitment with complete and enforceable contracts is available, equilibrium will be efficient. Capital and labor will receive their *ex ante* equilibrium opportunity cost in joint production. An interior solution in this case requires that entry condition (1) holds with equality. Denoting efficient levels by asterisks, and recalling that $w_k^* = 1$, the opportunity cost of labor is

$$w_\ell^* = y - 1.$$

Taking labor supply (7) into account, the efficient joint-production level of employment is

$$L^* = \left(\frac{y - 1}{2}\right)^\eta . \tag{8}$$

Finally, since labor receives its *ex ante* opportunity cost in joint production (i.e., $w_\ell^{j*} = w_\ell^*$), exit rule (6) in the efficient economy is simply

$$x < w_\ell^* . \tag{9}$$

B. *Incomplete-Contracts Equilibrium*

We now characterize equilibrium for an economy with no contractual precommitments, and compare it with the efficient economy.

Depressed Creation.—In general equilibrium, and off corners, entry condition (5) must hold with equality. As before, noting that $w_k = 1$ and taking labor supply (7) into account, the opportunity cost of labor and the level of joint-production employment are equal to

$$(10) \qquad w_\ell = y - 2 < w_\ell^*$$

$$(11) \qquad L = \left(\frac{y-2}{2}\right)^\eta < L^*.$$

With incomplete contracts, the "appropriating" factor's opportunity cost and the creation of joint-production units are depressed in general equilibrium. In order to ensure a proper return to the "appropriated" factor, the economic system must depress the opportunity cost of labor so that, capital receives its *ex ante* required return. This is done by underutilizing labor and, through this mechanism, weakening the bargaining position of "insider" labor.

Market Segmentation and Unemployment.—In equilibrium, the economy must generate a positive joint-production surplus s to compensate capital *ex post* for its *ex ante* opportunity cost [see equation (3)]. This implies that joint-production workers, who share in that surplus, must earn more than those who remain in autarky [see equation (4)]. The labor market is therefore "segmented" between those two sectors, and workers who remain in autarky are, in this sense, involuntarily "unemployed."

In fact, because capital was assumed to be infinitely elastic, the full cost of inefficiency is borne by the unemployed. The autarky "wage" w_ℓ is depressed in equilibrium so as to reduce the opportunity cost of labor and help compensate capital for appropriation by insider labor. The latter, on the other hand, suffer no welfare loss in the joint-production sector, and receive

$$(12) \qquad w_\ell^j = w_\ell + \frac{s}{2} = y - 1 = w_\ell^*.$$

Technological Sclerosis.—We now turn to the exit decision of a preexisting production unit with gross revenue x. From exit rule (6) in the microeconomics section, we know that the unit will exit if and only if its revenues are lower than $pw_\ell^j + (1-p)w_\ell$. It is easy to see that

$$pw_\ell^j + (1-p)w_\ell < w_\ell^*$$

since, by (10) and (12), we know that $w_\ell < w_\ell^*$ and $w_\ell^j = w_\ell^*$, and the probability of finding joint-production employment is $p = L < 1$. Comparing (6) and (9), this implies that the incomplete-contracts economy keeps in operation low-productivity units that would be scrapped in an efficient world. Not only is creation depressed by the appropriability problem, but also pressure on outdated units is reduced by the equilibrium decline of the opportunity cost of labor. This combination makes for a powerful drag on growth.

Scrambling.—Efficiently, the scrapping rule is based on a strict productivity ranking of preexisting production units, with the least productive units exiting, and others surviving. The ranking that forms the basis for the scrapping rule could be "scrambled" in the presence of appropriability and bargaining inefficiencies. Suppose we had introduced private inefficiency in separations of some form. Some high-productivity units could be destroyed because of private inefficiency, even though they would survive in the efficient economy; other, much less productive units, if they suffer less from private inefficiency, may survive because of "sclerosis," even though they would be destroyed in the efficient economy.

C. *Macroeconomic Response to Shocks*

Specificity together with incomplete contracting not only affects levels, but also the cyclical features of the economy. We now turn to the implications of appropriability for the economy's response to aggregate shocks.

Excessive Destruction.—Paradoxically, although scrapping is insufficient when compared with the efficient economy, it is *excessive* given the level of creation. To see this, note that the *social* shadow wage in this economy is w_ℓ, while the shadow wage used in the *private* destruction condition (6) is $pw_\ell^j + (1-p)w_\ell > w_\ell$ (since $w_\ell^j > w_\ell$).

Excessive destruction captures an important aspect of the employment crises that are feared during major macroeconomic adjustment episodes (see Caballero and Hammour, 1996b). The coexistence of sclerosis with excessive destruction also uncovers the fallacious nature, in our context, of the "liquidationist" argument, which sees in a contraction a way to improve welfare by reducing sclerosis.

Elastification.—Consider the economy's response to an aggregate shock to gross revenues y. Differentiating the equilibrium employment expressions (8) and (11), one can easily show that

$$\frac{dL/L}{dy/y} > \frac{dL^*/L^*}{dy/y}.$$

In other words, the cyclical response of the incomplete-contracts economy is more elastic than that of the efficient economy.[1] A decline in profitability requires a decline in the cost of labor to ensure capital of its required return. This is more difficult to achieve in the presence of appropriable quasi-rents and requires a proportionally larger decline in employment. More generally, in Caballero and Hammour (1996c) we show that, when the unemployed factor is the less elastic of the two, its cyclical responsiveness will inherit some of the elasticity of the other factor in the presence of appropriability, a phenomenon we term "elastification."

Decoupling.—In terms of the joint behavior of creation and destruction flows over the cycle, excessive destruction leads to a "decoupling" of creation and destruction, characterized by a destruction of jobs that is not accompanied by a simultaneous surge in job creation. In Caballero and Hammour (1996a), we argue that this manifestation of inefficiency is broadly consistent with what we know about job flows; in Caballero and Hammour (1996b), we argue for its relevance in accounting for massive employment and output declines during major adjustment episodes.

[1] If $3 < A < 4$, the elastification result is more stark. The efficient economy exhibits full joint-production employment in that case, while the inefficient economy does not.

III. Policy Considerations

At a very general level, the multidimensional macroeconomic problem we have described in the previous section stems from ill-defined property rights. These have long been of central concern to development economists, whether they affect the employment relation stressed here, or others, such as external finance. The design of institutions with the ability to safeguard proper contract enforcement is a main ingredient of a successful development strategy, an old idea that has acquired heightened relevance in the recent Eastern European reform episodes.

However, there is no need to go to such extreme episodes. The disruptive consequences of ill-defined property rights are present even in economies with the most developed judicial and regulatory systems. There are aspects of relationships that are simply too difficult to regulate contractually and to verify so as to fully protect the investing parties. Under those circumstances, and as a complement to institutional reforms, there is a need to design aggregate policies intended expressly to remedy the *macroeconomic* consequences of entrenched appropriability problems.

The first and most direct impact of appropriability is underinvestment. Thus a key element in an effective policy package is some sort of creation incentive. But, in isolation, the benefits of such a policy may come at a significant cost. Because of the wage pressures of increased hiring, a creation incentive may exacerbate the excessive-destruction problem. Thus, excessive destruction calls for another type of policy aimed at protecting existing production units. But, in isolation, such a policy does not address the need for more creation incentives and, depending on the form it takes, may actually further depress creation.

The problem of these two approaches considered in isolation is that, while partially correcting a problem along one margin, they fail to address—and often exacerbate—the problem on the other. A well-designed policy package must not only raise the level of creation, but must also succeed in "synchronizing" de-

struction with it. Only a well-balanced combination of creation incentives and protection of the existing structure can succeed in achieving those goals. In fact, it is straightforward to show in our simple model that both problems are solved by *simultaneously* implementing an investment subsidy equal to w_k, and a job-protection subsidy $L^*s^*/2$ to preexisting units if they remain in operation.

This two-pronged approach offers a useful framework to discuss macroeconomic policy in a variety of contexts. On adjustment strategies, it highlights the limitations of the "gradualism versus cold-turkey" debate, which, in its purest form, is a debate confined within the set of single-instrument policies aimed at slowing destruction. In Caballero and Hammour (1996b), we argue that Korea provides a good example of the managed-adjustment policies that stem from our framework, where investment incentives through the credit market were complemented with protection of domestic sectors, as well as with a direct intervention in labor markets to keep wage demands commensurate with productivity gains. Over the business cycle, our approach speaks on the soundness of countercyclical policies aimed simultaneously at encouraging investment and protecting employment during downturns.

REFERENCES

Caballero, Ricardo J. and Hammour, Mohamad L. "The Cleansing Effect of Recessions." *American Economic Review*, December 1994, *84*(5), pp. 1350–68.

_____. "On the Timing and Efficiency of Creative Destruction." *Quarterly Journal of Economics*, 1996a (forthcoming).

_____. "On the Ills of Adjustment." *Journal of Development Economics*, 1996b (forthcoming).

_____. "The Macroeconomics of Specificity." Mimeo, Massachusetts Institute of Technology, March 1996c.

Grout, Paul A. "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach." *Econometrica*, March 1984, *52*(2), pp. 449–60.

Hall, Robert E. and Lazear, Edward P. "The Excess Sensitivity of Layoffs and Quits to Demand." *Journal of Labor Economics*, April 1984, *2*(2), pp. 233–57.

Hart, Oliver and Moore, John. "A Theory of Debt Based on the Inalienability of Human Capital." *Quarterly Journal of Economics*, November 1994, *109*(4), pp. 841–80.

Klein, Benjamin; Crawford, Robert G. and Alchian, Armen A. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." *Journal of Law and Economics*, October 1978, *21*(2), pp. 297–326.

Myerson, Roger B. and Satterthwaite, Mark A. "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory*, April 1983, *29*(2), pp. 265–81.

Williamson, Oliver E. "Transaction-Cost Economics: The Governance of Contractual Relations." *Journal of Law and Economics*, October 1979, *22*(2), pp. 233–61.