Development Economics

By Debraj Ray, New York University


1 Introduction

What we know as the developing world is approximately the group of countries classified by the World Bank as having “low” and “middle” income. An exact description is unnecessary and not too revealing; suffice it to observe that these countries make up over 5 billion of world population, leaving out the approximately one billion who are part of the “high” income developed world. Together, the low and middle income countries generate approximately 6 trillion (2001) dollars of national income, to be contrasted with the 25 trillion generated by high income countries. An index of income that controls for purchasing power would place these latter numbers far closer together (approximately 20 trillion and 26 trillion, according to the World Development Report (2003)) but the per-capita disparities are large and obvious, and to those encountering them for the first time, still extraordinary.

Development Economics, a subject that studies the economics of the developing world, has made excellent use of economic theory, econometric methods, sociology, anthropology, political science, biology and demography and has burgeoned into one of the liveliest areas of research in all the social sciences. My limited approach in this brief article is one of deliberate selection of a few conceptual points that I consider to be central to our thinking about the subject. The reader interested in a more comprehensive overview is advised to look elsewhere (for example, at Dasgupta (1993), Hoff, Braverman and Stiglitz (1993), Ray (1998), Bardhan and Udry (1999), Mookherjee and Ray (2001), and Sen (1999)).

I begin with a traditional framework of development, one defined by conventional growth theory. This approach develops the hypothesis that given certain parameters, say savings or fertility rates, economies inevitably move towards some steady state. If these parameters are the same across economies, then in the long run all economies converge to one another. If in reality we see utter lack of such convergence — which we do (see, e.g., Quah (1996) and Pritchett (1997)) — then such an absence must be traced to a presumption that the parameters in question are not the same. To the extent that history plays any role at all in this view, it does so by affecting these parameters — savings, demographics, government interventionism, “corruption” or “culture”.
This view is problematic for reasons that I attempt to clarify below. Indeed, the bulk of my essay is organized around the opposite presumption: that two societies with the same fundamentals can evolve along very different lines — going forward — depending on past expectations, aspirations or actual history.

Now, after a point, the distinction between evolution and parameter is a semantic one. By throwing enough state variables (“parameters”) into the mix, one might argue that there is no difference at all between the two approaches. Formally, that would be correct, but then “parameters” would have to be interpreted broadly enough so as to be of little explanatory value. Ahistorical convergence and historically conditioned divergence express two fundamentally different world views, and there is little that semantic jugglery can do to bring them together.

2 Development From The Viewpoint of Convergence

Why are some countries poor while others are rich? What explains the success stories of economic development, and how can we learn from the failures? How do we make sense of the enormous inequalities that we see, both within and across questions? These, among others, are the “big questions” of economic development.

It is fair to say that the model of economic growth pioneered by Robert Solow (1956) has had a fundamental impact on “big-question” development economics. For theory, calibration and empirical exercises that begin from this starting point, see, e.g., Lucas (1990), Mankiw, Romer and Weil (1992), Barro (1991), Parente and Prescott (2000) and Banerjee and Duflo (2005). Solow’s pathbreaking work introduced the notion of convergence: countries with a low endowment of capital relative to labor will have a high rate of return to capital (by the “law” of diminishing returns). Consequently, a given addition to the capital stock will have a larger impact on per-capita income. It follows that, controlling for parameters such as savings rates and population growth rates, poorer countries will tend to grow faster and hence will catch up, converge to the levels of well-being enjoyed by their richer counterparts. Under this view, development is largely a matter of getting some economic and demographic parameters right and then settling down to wait.

To be sure, savings and demography are not the only factors that qualify the argument. Anything that systematically affects the marginal addition to per-capita income must be controlled for, including variables such as investment in “human capital” or harder-to-quantify factors such as “political climate” or “corruption”. A failure to observe convergence must be traced to one or another of these “parameters”.

Convergence relies on diminishing returns to “capital”. If this is our assumed starting
point, the share of capital in national income does give us rough estimates of the concavity of production in capital. The problem is that the resulting concavity understates observed variation in cross-country income by orders of magnitude. For instance, Parente and Prescott (2000) calibrate a basic Cobb-Douglas production function by using reasonable estimates of the share of capital income (0.25), but then huge variations in the savings rate do not change world income by much. For instance, doubling the savings rate leads to a change in steady state income by a factor of 1.25, which is inadequate to explain an observed range of around 20:1 (PPP). Indeed, as Lucas (1990) observes, the discrepancy actually appears in a more primitive way, at the level of the production function. For the same simple production function to fit the data on per-capita income differences, a poor country would have to have enormously higher rates of return to capital; say, 60 times higher if it is one-fifteenth as rich. This is implausible. And so begins the hunt for other factors that might explain the difference. What did we not control for, but should have?

This describes the methodological approach. The convergence benchmark must be pitted against the empirical evidence on world income distributions, savings rates, or rates of return to capital. The two will usually fail to agree. Then we look for the parametric differences that will bridge the model to the data.

“Human capital” is often used as a first port of call: might differences here account for observed cross-country variation? The easiest way to slip differences in human capital into the Solow equations is to renormalize labor. Usually, this exercise does not take us very far. Depending on whether we conduct the Lucas exercise or the Prescott-Parente variant, we would still be predicting that the rate of return to capital is far higher in India than in the U.S., or that per-capita income differences are only around half as much (or less) as they truly are. The rest must be attributed to that familiar black box: “technological differences”. That slot can be filled in a variety of ways: externalities arising from human capital, incomplete diffusion of technology, excessive government intervention, within-country misallocation of resources, .... All of these — and more — are interesting candidates, but by now we have wandered far from the original convergence model, and if at all that model still continues to illuminate, it is by way of occasional return to the recalibration exercise, after choosing plausible specifications for each of these potential explanations.

This model serves as a quick and ready fix on the world, and it organizes a search for possible explanations. Taken with the appropriate quantity of salt, and viewed as a first pass, such an exercise can be immensely useful. Yet playing this game too seriously reveals a particular world-view. It suggests a fundamental belief that the world economy is ultimately a great leveller, and that if the levelling is not taking place we must search for that explanation in parameters that are somehow structurally rooted in a society.
To be sure, the parameters identified in these calibration exercises do go hand in hand with underdevelopment. So do bad nutrition, high mortality rates, or lack of access to sanitation, safe water and housing. Yet there is no ultimate causal chain: many of these features go hand in hand with low income in self-reinforcing interplay. By the same token, corruption, culture, procreation and politics are all up for serious cross-examination: just because “cultural factors” (for instance) seems more weighty an “explanation” does not permit us to assign it the status of a truly exogenous variable.

In other words, the convergence predicted by technologically diminishing returns to inputs should not blind us to the possibility of nonconvergent behavior when all variables are treated as they should be — as variables that potentially make for underdevelopment, but also as variables that are profoundly affected by the development process.

3 Development from The Viewpoint of Nonconvergence

This leads to a different way of asking the big questions, one that is not grounded in any presumption of convergence. The starting point is that two economies with the same fundamentals can move apart along very different paths. Some of the best-known economists writing on development in the first half of the twentieth century were instinctively drawn to this view: Young (1928), Nurkse (1953), Leibenstein (1957) and Myrdal (1957) among them.

Historical legacies need not be limited to a nation’s inheritance of capital stock or GDP from its ancestors. Factors as diverse as the distribution of economic or political power, legal structure, traditions, group reputations, colonial heritage and specific institutional settings may serve as initial conditions — with a long reach. Even the accumulated baggage of unfulfilled aspirations or depressed expectations may echo into the future. Factors that have received special attention in the literature include historical inequalities, the nature of colonial settlement, the character of early industry and agriculture, and early political institutions.

3.1 Expectations and Development

Consider the role of expectations. Rosenstein-Rodan (1943) and Hirschman (1958) (and several others following them) argued that economic development could be thought of as a massive coordination failure, in which several investments do not occur simply because other complementary investments are similarly depressed in the same bootstrapped way. Thus one might conceive of two (or more) equilibria under the very same fundamental conditions, “ranked” by different levels of investment.
Such “ranked equilibria” reply on the presence of a complementarity: a particular form of externality in which the taking of an action by an agent increases the marginal benefit to other agents from taking the a similar action. In the argument above, sector-specific investments lie at the heart of the complementarity: more investment in one sector raises the return to investment in some related sector.

Once complementarities — and their implications for equilibrium multiplicity — enter our way of thinking, they seem to pop up everywhere. Complementarities play a role in explaining how technological inefficiencies persist (David (1985), Arthur (1994)), why financial depth is low (and growth volatile) in developing countries (Acemoglu and Zilibotti (1997)), how investments in physical and human capital may be depressed (Romer (1986), Lucas (1988)), why corruption may be self-sustaining (Kingston (2005), Emerson (2006)), the growth of cities (Henderson (1988), Krugman (1991)), the suddenness of currency crises (Obstfeld (1994)), or the fertility transition (Munshi and Myaux (2006)); I could easily go on. Even the traditional Rosenstein-Rodan view of demand complementarities has been formally resurrected (Murphy, Shleifer and Vishny (1989)).

An important problem with theories of multiple equilibrium is that they carry an unclear burden of history. Suppose, for instance, that an economy has been in a low-level investment trap for decades. Nothing in the theory prevents that very same economy from abruptly shooting into the high-level equilibrium today. There is a literature that studies how the past might weigh on the present when a multiple equilibrium model is embedded in real time (see, e.g., Adserà and Ray (1998) and Frankel and Pauzner (2000)). When we have a better knowledge of such models we will be able to make more sense of some classical issues, such as the debate on balanced versus unbalanced growth. Rosenstein-Rodan argued that a “big push” — a large, balanced infusion of funds — is ideal for catapulting an economy away from a low-level equilibrium trap. Hirschman argued, in contrast, that certain “leading sectors” should be given all the attention, the resulting imbalance in the economy provoking salubrious cycles of private investment in the complementary sectors. To my knowledge, we still lack good theories to examine such debates in a satisfactory way.

3.2 Aspirations, Mindsets and Development

The aspirations of a society are conditioned by its circumstances and history, but they also determine its future. There is scope, then, for a self-sustaining failure of aspirations and economic outcomes, just as there is for ever-progressive growth in them (Appadurai (2004), Ray (2006)).

Typically, the aspirations of an individual are generated and conditioned by the experiences of others in her “cognitive neighborhood”. There may be several reasons for this:
the use of role models, the importance of relative income, the transmission of information, or peer-determined setting of internal standards and goals. Such conditioning will affect numerous important socio-economic outcomes: the rate of savings, the decision to migrate, fertility choices, technology adoption, the adherence to norms, the choice of ethnic or religious identity, the work ethic, or the strength of mutual insurance motives.

As an illustration, consider the notion of an aspirations gap. In a relatively narrow economic context (though there is no need to restrict oneself to this) such a gap is simply the difference between the standard of living that’s aspired to and the standard of living that one already has. The former isn’t exogenous; it will depend on the ambient standards of living among peers or near-peers, or perhaps other communities.

The aspirations gap may be filled — or neglected — by deliberate action. Investments in education, health, or income-generating activities are obvious examples. Does history, via the creation of aspirations gaps, harden existing inequalities and generate poverty traps? Or does the existence of a gap spur individuals on ever harder to narrow the distance? As I have argued in Ray (1998, Sections 3.3.2 and 7.2.4) and Ray (2006), the effect could go either way. A small gap may encourage investments, a large gap stifle it. This leads not only to history-dependence, but also a potential theory of the connections between income inequality and the rate of growth.

These remarks are related to Duflo’s (2006) more general (but less structured) hypothesis that “being poor almost certainly affects the way people think and decide”. This “mindset effect” can manifest itself in many ways (an aspirations gap being just one of them), and can lead to poverty traps. For instance, Duflo and Udry (2004) find that certain within-family insurance opportunities seem to be inexplicably foregone. In broadly similar vein, Udry (1996) finds that men and women in the same household farm land in a way that is not Pareto-efficient (gains in efficiency are to be had by simply reallocating inputs to the women’s plots). These observations suggest a theory of the poor household in which different sources of income are treated differently by members of the household, perhaps in the fear that this will affect threat points in some intrahousehold bargaining game. This in itself is perhaps not unusual, but the evidence suggests that poverty itself heightens the salience of such a framework.

3.3 Markets and History-Dependence

I now move on to other pathways for history-dependence, beginning with the central role of inequality. According to this view, historic inequalities persist (or widen) because each individual entity — dynasty, region, country — is swept along in a self-perpetuating path of occupational choice, income, consumption, and accumulation. The relatively poor may be limited in their ability to invest productively, both in themselves and in their children.
Such investments might include both physical projects such as starting a business, or “human projects” such as nutrition, health and education. Or the poor may have ideas that they cannot profitably implement, because implementation requires startup funds that they do not have. Yet, faced with a different level of initial inequality, or jolted by a one-time redistribution, the very same economy may perform very differently. Now investment opportunities are available widely through the population, and a new outcome emerges with not just lower inequality, but higher aggregate income. These are different steady states, and they could well be driven by distant histories (see, e.g., Dasgupta and Ray (1986), Banerjee and Newman (1993), Galor and Zeira (1993), Ljungqvist (1993), Ray and Streufert (1993), Piketty (1997) or Matsuyama (2000)).

The intelligent layperson would be unimpressed by the originality of this argument. That the past systematically preys on the present is hardly rocket science. Yet theories based on convergence would rule such obvious arguments out. Under convergence, the very fact that the poor have limited capital relative to labor allows them to grow faster and (ultimately) to catch up. Economists are so used to the convergence mechanism that they sometimes do not appreciate just how unintuitive it is.

That said, it is time now to cross-examine our intelligent layperson. For instance, if all individuals have access to a well-functioning capital market, they should be able to make an efficient economic choice with no heed to their starting position, and the shadows cast by past inequalities must disappear (or at least dramatically shrink). For past wealth to alter current investments, imperfections in capital or insurance markets must play a central role.

At the same time, such imperfections aren’t sufficient: the concavity of investment returns would still guarantee convergence. A first response is that such “production functions” are simply not concave. A variety of investment activities have substantial fixed costs: business startups, nutritional or health investments, educational choices, migration decisions, crop adoptions. Indeed, it is hard to see how the presence of such nonconvexities could not be salient for the ultrapoor. Coupled with missing capital markets, it is easy to see that steady state traps, in which poverty breeds poverty, are a natural outcome (see, e.g., Majumdar and Mitra (1982), Galor and Zeira (1993)). Surveys of the economic conditions of the poor (Banerjee and Duflo (2007), Fields (1980)) are eminently consistent with this point of view.

A related source of nonconvexity arises from limited liability. A highly indebted economic agent may have little incentive to invest. Similarly, poor agents may enter into contracts with explicit or implicit lower bounds on liability. These bounds can create poverty traps (Mookherjee and Ray (2002a)).

Investment activities that go past these minimal thresholds are potentially open to “convexification”. There are various stopping points for human capital acquisition, and a
household can hold financial assets which are, in the end, scaled-down claims on other businesses. Under this point of view, dynasties that make it past the ultrapoor thresholds will exhibit ergodic behavior (as in Loury (1981) and Becker and Tomes (1986)) and so the prediction is roughly that of a two-class society: the ultrapoor caught in a poverty trap and the remainder enjoying the benefits of convergence. History would matter in determining the steady-state proportions of the ultrapoor.

But this sort of analysis ignores the endogenous nonconvexities brought about by the price system. For instance, even if there are many different education levels, the wage payoff to such level will generally be determined by the market. There is good reason to argue (see, e.g., Ljungqvist (1993), Freeman (1996) and Mookherjee and Ray (2002b, 2003)) that the price system will sort individuals into different occupational choices, and that there will be persistent inequality across dynasties located at each of these occupational slots. Thus an augmented theory of history dependence might predict a particular proportion of the ultrapoor trapped by physical nonconvexities (low nutrition, ill-health, debt, lack of access to primary education), as well as a persistently unequal dispersion of dynasties across different occupational choices, induced by the pecuniary externalities of relative prices.

Note that it is precisely the high-inequality, high-poverty steady states that are correlated with low average incomes for society as a whole, and it is certainly possible to build a view of underdevelopment from this basic premise. The argument can be bolstered by consideration of economy-wide externalities; for instance, in physical and human capital (Romer (1986), Lucas (1988), Azariadis and Drazen (1990)).

3.4 History, Aggregates and the Interactive World

Theories such as these might yield a useful model for the interactive world economy. Take, for instance, the notion of aspirations. Just as domestic aspirations drive the dynamics of accumulation within countries, there is a role, too, for national aspirations, driven by inter-country disparities in consumption and wealth, and its effect on the international distribution of income. Even the simplest growth framework that exhibits the usual features of convexity in its technology and budget constraints could give rise in the end to a world distribution that is bipolar. Countries in the middle of that distribution would tend to accumulate faster, be more dynamic and take more risks as they see the possibility of full catch-up within a generation or less. One might expect the greatest degree of “country mobility” in this range. In contrast, societies that are far away from the economic frontier may see economic growth as too limited and too long-term an instrument, leading to a failure, as it were, of “international aspirations”. Groups within these societies may well resort to other methods of potential economic
gain, such as rent-seeking or conflict. (The aggregate impact of such activities would reinforce the slide, of course.)

Of course, an entirely mechanical transplantation of the aspirations model to an international context isn’t a good idea. Countries are not individual units: a more complete theory must take into account the aspirations of various groups in the different countries, and the domestic and international components that drive such aspirations.

Next, consider the role of markets. Once again, tentatively view each country as a single economic agent in the framework of Section 3.3. Now the nonconvexities to be considered are at the level of the country as a whole — Young’s increasing returns on a grand scale, or economy-wide externalities as in Lucas-Azariadis-Drazen. This reinterpretation is fairly standard, but less obviously, the occupational choice story bears reinterpretation as well. To see this, note that the pattern of production and trade in the world economy will be driven by patterns of comparative advantage across countries. But in a dynamic framework, barring nonreproducible resources such as land or mineral endowments, every endowment is potentially accumulable, so that comparative advantage becomes endogenous. Thus we may view countries as settling into subsets of occupational slots (broadly conceived), producing an incomplete range of goods and services relative to the world list, and engaging in trade.

For instance, suppose that country-level infrastructure is suitable for either high-tech or low-tech production, but not both. If both high-tech and low-tech are important in world production and consumption, then some country has to focus on low-tech and another on high-tech. Initial history will constrain such choices, if for no reason than the fact that existing infrastructure (and national wealth) determines the selection of future infrastructure. This is not to say that no country can break free of those shackles. For instance, as the whole world climbs up the income scale, natural nonhomotheticities in demand will push composition more and more in favor of high-quality goods. As this happens, more and more countries will be able to make the transition. But on the whole, if national infrastructure is more or less conducive to some (but not the full) range of goods, the nonconvergence model that we discussed for the domestic economy must apply to the world economy as well.

This raises an obvious question. What is so specific about “national infrastructure”? Why is it not possible for the world to ultimately rearrange itself so that every country produces the same or similar mix of goods, thus guaranteeing convergence? Do current national advantages somehow manifest themselves in future advantages as well, thus ensuring that the world economy settles into a permanent state of global inequality? Might economic underdevelopment across countries, at least in this relative sense, always stay with us?

To properly address such questions we have to drop the tentative assumption that each
country can be viewed as an individual unit. In a more general setting, there are individuals within countries, and then there is cross-country interaction. The former are subject to the forces of occupational structure (and possible fixed costs), as discussed in Section 3.3. The latter are subject to the specificities, if any, of “national infrastructure”, determining whether countries as a whole have to specialize, at least to some degree. The relative importance of within-country versus cross-country inequalities will rest, in large part, on considerations such as these.

I haven’t brought in international political economy so far (though see below). Yet, as frameworks go, this is not a bad one to start thinking about the effects of globalization. It is certainly preferable to a view of the world as a set of disconnected, autarkic growth models.

### 3.5 Institutions and History

In many developing countries, the early institutions of colonial rule were directly set up for the purposes of surplus extraction. There would be variation, of course, depending on whether the areas were sparsely or densely populated to begin with, or whether there was large-scale availability of mineral deposits. Resource deposits certainly favored large-scale extractive industry (as in parts of South America), while soil and weather conditions might encourage plantation agriculture, often with the use of slave labor (as in the Caribbean). On the other hand, a high preexisting population density would favor extraction of a different hue: the setting-up of institutional systems to acquire rents (the British colonial approach in large parts of India).

It has been argued, perhaps most eloquently by Sokoloff and Engerman (2000), that initial institutional modes of production and extraction in distant history had far-reaching effects on subsequent development. In their words, scholars “have begun to explore the possibility that initial conditions, or factor endowments broadly conceived, could have had profound and enduring impacts on long-run paths of institutional and economic development . . . ”. Such inequalities may then be inimical to development in a variety of ways (such as the market based pathways discussed earlier). In contrast, where initial settlements did not go hand in hand with systems of tribute, land grants, or large-scale extractive industries (as in several regions of North America), one might expect comparative equality and a subsequent path of development that is more broad-based.

This is consistent with the market-based processes considered earlier. But a principal strand of the Sokoloff-Engerman argument, as also the lines of reasoning pursued in Robinson (1998), Acemoglu, Johnson and Robinson (2001, 2002) and Acemoglu (2006), emphasizes political economy. In the words of Sokoloff and Engerman (2000), “initial conditions had lingering effects . . . because government policies and other institutions
tended to reproduce them. Specifically, in those societies that began with extreme inequality, elites were better able to establish a legal framework that insured them disproportionate shares of political power, and to use that greater influence to establish rules, laws, and other government policies that advantaged members of the elite relative to nonmembers contributing to persistence over time of the high degree of inequality . . . In societies that began with greater equality or homogeneity among the population, however, efforts by elites to institutionalize an unequal distribution of political power were relatively unsuccessful . . .

The elite —erswhile collectors of tribute, land-grant recipients, plantation owners and the like — may survive long after the initial institutions that spawned them are gone. Such survival may nevertheless be quite compatible with the maximization of aggregate surplus provided that the elite are the most efficient of the economic citizenry in the generations to come. But of course, there is absolutely no reason why this should be the case. A new generation of entrepreneurs, economic and political, may be waiting to take over in the wings. It is an open question as to what will happen next, but often, the elite may well engage in policy that has its goal not economic efficiency but the crippling of political opposition. Some evidence of this reluctance to let go may be seen in literature that argues that more unequal societies redistribute less (see Perotti (1994, 1996), and the survey by Bénabou (1996)).

There are other routes. The elite may be unable to avoid an oppositional showdown. A theory of bad policy may then have to be replaced by model of social unrest and conflict generated by initial inequality. While this mechanism is clearly different, the end result is the same. The channeling of resources to ongoing conflict will surely inhibit the accumulation of productive resources (Benhabib and Rustichini (1996), González (2007)). There may also be effects running through legal systems (see, e.g., La Porta, Lopez-de-Silanes, Shleifer and Vishny, (1997, 1998)) or the varying nature of different colonial systems (see, e.g., Bertocchi and Canova (2002)). There may be effects running through the insecurity of property rights of fear of elite expropriation (see, e.g.,Binswanger, Deininger and Feder (1995)).

We do not yet have a systematic exploration of these mechanisms, nor an accounting of their relative importance. But there is some reduced-form evidence that historical institutions do affect growth in the manner described by Sokoloff and Engerman. The problem in establishing an empirical assertion of this sort is fairly obvious: good institutions and good economic outcomes may simply be correlated via variables we fail to observe or measure, or any observed causality may simply run from outcomes to institutions. Acemoglu, Johnson, and Robinson (2001) propose a novel instrument for (bad) institutions: the mortality rate among European settlers (bishops, sailors and soldiers to be exact). This is a clever idea that exploits the following theory: only areas that could be settled by the Europeans developed egalitarian, broadbased institutions. In the other
areas, the same Europeans settled for slavery, dictatorship, highly unequal land grants and unbridled extraction instead. (The implied instrument is more convincing when the analysis is combined with controls for the general disease environment, which could have a direct effect on performance)

The Acemoglu-Johnson-Robinson results, which show that early institutions have an effect on current performance, are provocative and interesting. It bears reiteration, though, that IV estimates are suggestive of an institutional impact on development, but one just cannot be sure of what the mechanism is. By relinquishing more immediate institutional effects on the grounds of, say, endogeneity, it becomes that much harder to figure out the structural pathways of influence. This appears to be an endemic problem with large, sweeping cross-country studies that attempt to detect an institutional effect. Good instruments are hard to find, and when they exist, their effect could be the echo of one or more of a diversity of underlying mechanisms.

Iyer (2004) and Banerjee and Iyer (2005) consider a somewhat different channel of influence. Both these paper study the differential impact of colonial rule within a single country, India. Iyer studies British annexations of parts of India, and the effect today on public goods provision across annexed and non-annexed parts. There is obvious endogeneity in the areas chosen for annexation (a similar observation applies, in passing, to countries “selected” for colonization). Iyer instruments annexation by exploiting the so-called Doctrine of Lapse, under which the British annexed states in which a native ruler died without a biological heir. Banerjee and Iyer study the effect of variations in the land revenue systems set up by the British, starting from the latter half of the eighteenth century. In particular, they distinguish between landlord-based institutions, in which large landlords were used to siphon surplus to the British, and other areas based on rent payments, either directly from the cultivator or via village bodies. While these institutions of extraction no longer exist (India has no agricultural income tax), the authors argue that divided, unequal areas in the past cannot come together for collective action. Dispossessed groups are more worried about insecurity of tenure and fear of expropriation than about the absence of public goods, investment (public or private) or development expenditure.

### 3.6 Institutions and the Interactive World

In Section 3.4, we applied market-based theories of occupational choice and persistent inequality to the interactive world economy, (tentatively) treating each country as an economic agent. Recall the main assumption for such an interpretation to be sensible: that countries must face infrastructural constraints that limit full diversification. With these constraints in place, there will be persistent inequality in the world income distribution, with countries in “occupational niches” that correspond to their infrastructural
choices.

Bring to this story the role of institutional origins. Then a particular institutional history may be more suited to particular subsets of occupations, driving the country in question into a determinate slot in the world economy. From that point on, the persistent cross-country inequalities generated by the market-based theory will continue to link past institutions to subsequent growth. In short, initial institutional differences may be correlated with subsequent performance, but the magnitude of that under- or over-performance is not to be entirely traced to initial history. Distant history could simply have served as a marker for some countries to supply a particular range of occupations, goods and services. Today’s inequality may well be driven, not by that far-away history but simply by the world equilibrium path that follows on those initial conditions. If all goods are needed, there must be banana producers, sugar manufacturers, coffee growers, and high-tech enclaves, but there cannot be too little or too many of any of them.

The “inefficient political power” argument used in Section 3.5 can also be transplanted to international interactions. It may well be that a large part of such interactions — protection of international property rights, restrictions on technology transfer, or barriers to trade — is used to deter the entry of developing countries onto a level playing field in which they can successfully compete with their compatriots in developed countries. It would certainly be naive to disregard this point of view altogether.

Looked at this way, our view of history fits in well with the entire debate on globalization. One might view one side of this debate as emphasizing the convergence attributes of globalization: outsourcing, the establishment of international production standards, technology transfer, political accountability, responsible macroeconomic policies may all be invoked as footsoldiers in the service of convergence.

On the other side of the battlelines are equally formidable opponents. A skewed playing field can only keep tipping, so goes the argument. The protection of intellectual property is just a way of maintaining or widening existing gaps in knowledge. Technology transfers are inappropriate because the input mix isn’t right. Nonconvexities and increasing returns are endemic.

My goal here isn’t to take sides on this debate (though like everyone, I do have an opinion) but to clarify it from a “nonconvergence perspective” that has so far received more attention within the closed economy. There is a strong parallel between globalization (and those contented or discontented with it, to borrow a phrase from Joseph Stiglitz (2002)) and the questions of convergence and divergence in closed economies.
4  Digging Deeper: The Microeconomics of Development

There is no getting away from the big questions, even if they cannot be fully answered with the knowledge and tools we have at hand. The issues we’ve discussed (and our intuitive first-takes on them) determine our world view, the cognitive canvas on which we arrange our overall thoughts. But only the most hard-bitten macroeconomist would feel no trepidation about taking these models literally, and applying them without hesitation across countries, regions and cultures.

The microeconomics of development enables us to dig below the macro questions, unearthing insight and structure with far more confidence than we can hope to have at the world or cross-country level. From the viewpoint of economic theory, the assumptions made can be more carefully motivated and are open to careful testing. From the viewpoint of empirical analysis, it is far easier to find instruments or natural experiments, or to conduct one’s own experiments for that matter. There is, no doubt, the philosophical problems of scaling the results up, of using a well-controlled finding to predict outcomes elsewhere. In the end, the choice between the fuzzy, imprecise big picture and the small yet carefully delineated canvas is perhaps a matter of taste.

I need hardly add that my selectivity continues unabated: there is an entire host of issues, and I can but touch on a fraction of them. I focus deliberately on four important topics that are both relevant to my overall theme of history-dependence, and have been the subject of much recent attention.

4.1 The Credit Market

As we’ve seen, a failure of the credit market to function is at the heart of market-based arguments for divergence.

The fundamental reason for imperfect or missing credit markets is that individuals cannot be counted upon (for reasons of strategy or luck) to fully repay their loans. If borrowers do not have deep pockets, or if a well-defined system to enforce repayment is missing, then it stands to reason that lenders would be reluctant to advance those loans in the first place. There is little point in asserting that a well-chosen risk-premium will deal with these risks: the premium itself affects the default probability. Therefore some borrowers will be shut out of the market, no matter what rate of interest they are willing to pay. Such a market will typically clear by rationing access to credit, and not by an adjustment of the rate of interest.

Three fundamental features characterize different theories of imperfect credit markets. There is classical adverse selection, in which borrower (or project) characteristics may systematically adjust with the terms of the loan contract on offer. Stiglitz and Weiss
(1981) initiate this literature for credit markets, arguing that the higher the interest rate, the more likely it is that the borrower pool will be contaminated by riskier types. Then there is the moral hazard problem (see, e.g., Aghion and Bolton (1997)), in which the borrower must expend effort *ex post* to increase the chances of project success. Moral hazard also ties into “debt overhang”, in which existing indebtedness makes it less credible that a borrower will put in sustained effort in the project. Finally, there is the enforcement problem (see, e.g., Eaton and Gersovitz (1981)), in which a borrower may be tempted to engage in strategic default. Ghosh, Mookherjee and Ray (2000) survey some of the literature.

The poor are particularly affected, not because they are intrinsically less trustworthy, but because in the event of a project failure, they will not have the deep pockets to pay up. The poor may well possess collateral — a small plot of land or their labor — but such collateral may be hard to adequately monetize. A formal-sector bank may be unwilling to accept a small rural plot as collateral, much less bonded labor! But other lenders (a rural landlord, for instance) might. It is therefore not surprising to see interlinkages in credit transactions for the poor: a small farmer is likely to borrow from a trader who trades his crop, while a rural tenant is likely to borrow from his landlord. Even when the entire market looks competitive, these niches may create pockets of exploitative local monopoly (Ray and Sengupta (1989), Floro and Yotopoulos (1991), Floro and Ray (1997), Mansuri (1997), Genicot (2000)).

In short, the very fact of their limited wealth puts the relatively poor under additional constraints in the credit market. This is why imperfect capital markets serve as a starting point for many of the models that study market-based history-dependence.

The direct empirical evidence on the existence of credit constraints is surprisingly sparse, which is obviously not to say that they don’t exist, but to point out that this is an area for future research. Existing literature in a development context largely uses the existence of (presumably undesirable) consumption fluctuations in households to infer the lack of perfect financial markets; see Morduch (1994), Townsend (1995) and Deaton (1997). A direct test for credit constraints yields positive results for Indian firms (Banerjee and Duflo (2004)), though it is unclear how general this finding is (see, e.g., Hurts and Lusardi (2004)). There is a sizeable literature which deals with the impact of credit constraints on outcomes such as health (Foster (1995)), education (Jacoby and Skoufias (1997)) or the acquisition of production inputs such as bullocks (Rosenzweig and Wolpin (1993)).

Chiappori and Salanie (2000) and Karlan and Zinman (2006) are two examples of specific tests for different frictions, such as adverse selection and enforcement. Udry’s seminal (1994) paper on credit and insurance markets in Northern Nigeria may be viewed as singling out enforcement as perhaps the most important binding constraint. The importance of enforcement constraints is, of course, not peculiar to credit or insurance;

Finally, there is a literature on microcredit, the lending of relatively small amounts to the very poor; Armendáriz and Morduch (2005) is a good starting point.

4.2 Collective Action for Public Goods

There is a growing literature on the political economy of development. Unlike some mainstream approaches in political science and political economy, this literature appears to largely eschew voting models. In my view this is not a bad thing. Perhaps the most important criticism of voting models is that even in vigorous democracies, most policies are not subject to referenda among the citizenry at large. Certainly, there are periodic elections, and the sum total of enacted policies — and the package of future promises — are then up for voter scrutiny, but nevertheless, there is a large and significant gap between voting and the enactment of a particular policy. Between that policy and the voter falls the shadow of collective action, lobbies, capture and influence, cynical tradeoffs across special interests, and covert or open conflict. For countries with a nondemocratic history, these considerations are expanded by orders of magnitude.

An important literature concerns the determinants of collective action for the provision of public goods, and how poverty or inequality affects the ability to engage in such action. The relationship here is complex. There are two potential reasons why inequality in a community may enhance collective action. First, the elite in a high-inequality community might largely internalize all the benefits from the resulting public good, and therefore pay for it (Olson (1965)). Good examples involve military alliances (Sandler and Forbes (1980)), technology adoption (Foster and Rosenzweig (1995)) or even “top-down interventions” by local rulers or elites (Banerjee, Iyer and Somanathan (2007)). Second, the elite have a low opportunity cost of money, while the poor have a low opportunity cost of labor; in some situations, the two resources can be usefully combined for collective action (an alliance for violent conflict, as in Esteban and Ray (2007a) is a good example). But there is a variety of situations in which inequality can dampen effective collective action: when all agents supply similar inputs — say effort — and their impact or cost of provision is nonlinear (Ray, Baland and Dagnielie (2006), Khwaja (2006)), when there are unequally distributed private endowments (Baland and Platteau (1998), Bardhan, Ghatak and Karaivanov (2006)), or when different individuals in the same community want different things by virtue of their social differences or inequality (Alesina, Baqir and Easterly (1999), Banerjee, Mookherjee, Munshi and Ray (2001),
Miguel and Gugerty (2005), Alesina and La Ferrara (2005)), or when inequalities in wealth erode the informational basis of collective action (Esteban and Ray (2006)). The importance of this area of research cannot be overemphasized. Several of the fundamental accompaniments of development requires state intervention at a basic level: health, education, social safety nets, infrastructure. This is especially so in poor countries, where privatized health and education are often ruled out by the sheer force of economic necessity. Yet states often are set upon by numerous claims that compete for their attention. How are these claims resolved? The theory and practice of collective action demands more research.

Moreover, while it can be argued (as I’ve done above) that inequality within a community might go either way in affecting that community’s ability to obtain public goods, there is no escaping the fact that at the level of the entire society, high inequality serves to fracture and divide. Simply put, the very rich want state policy that is different from the very poor, and rare is the society that has them in the same camp, and demanding the same things of their government. In the world of the median voter, one might simply resolve these issues by looking at the median voter’s ideal policy, but even in this rarefied scenario, there are complex issues that deserve our consideration. Political alliances can often redefine the median voter (Levy (2004)) and even without alliances it is unclear just who the median voter is (Bénabou (2000)). When we return to the “real world” of collective action, these issues are magnified considerably. For now each citizen does not have an endowment of one vote. The real endowments are labor and money. How these commodities combine (or compete) is fundamental to our understanding of political economy and — via this channel — our views on persistent history-dependence.

### 4.3 Conflict

A more sinister expression of collective action is conflict. In the second half of the twentieth century and well into the first decade of the twenty first, the loss of human life from conflicts in developing countries is immense; the costs are beyond measurement. Even the narrow economic costs of conflict can be extremely large (Hess (2003)).

That conflict contributes to economic regress is not surprising. But given our focus on history-dependence, it is of equal interest to consider the casual chain running from underdevelopment to conflict. That chain has a natural and simple foundation: poverty reduces the opportunity cost of engaging in conflict. The grabbing of resources, often in an organized way, is often a far more lucrative alternative to the steady process of wealth accumulation. It’s certainly quicker. (One might argue that there is less to gain as well, but this effect is attenuated in unequal societies.)

This unfortunate observation has substantial empirical support. For instance, Miguel,
Satyanath and Sergenti (2004) use rainfall as an instrument for economic growth in 41 African countries and derive a striking negative effect of growth on civil conflict: a negative growth shock of 5 percentage points raises the likelihood of civil conflict by 50%. See also Dube and Vargas (2007) and Hildalgo, Naidu, Nichter and Richardson (2007), both of which also instrument for economic shocks to find significant effects on conflictual outcomes. Collier and Hoeffler (1998), Sambanis (2001), Fearon and Laitin (2003), Do and Iyer (2006) all establish strong correlations between economic adversity and conflict, the last of these countries establishing this over regions in a single country (Nepal).

Yet conflict is demonstrably wasteful, and if warring parties could sit down at the negotiating table, why would societies engage in it? This is a classical question to which there are a number of possible answers. First, there may be a prisoner’s-dilemma-like quality to conflictual incidents, in the sense that one party can precipitate attacks while the other remains passive (Leventoglu and Slantchev (2005)). Second, while conflict generates waste, there is no reason to believe that every group is thereby made worse off by it. It is entirely possible that a group prefers conflict to a peaceful outcome: the former involves a smaller pie, but it also may involve a larger share of it (Esteban and Ray (2001)). Third, while one should be able to find a system of taxes and transfers that Pareto-dominate the conflict outcome, but for various reasons — lack of commitment, or a sparse informational base for the levying of taxes, dynamics with rapid power shifts — it may not be possible to implement that system (Fearon (1995), Powell (2004, 2006)). Fourth, it is certainly possible that conflict is over indivisible resources such as political power or religious hegemony. It may then be absurd to imagine that side A compensates side B with suitable transfers in exchange for political power: the lack of credibility involved is only too apparent. Finally, conflict may be endemic because both parties to it have incomplete information regarding chances of success, though this view has come under increasing criticism from political scientists (see, e.g., Fearon (1995)).

The next question of relevance concerns ethnic and social divisions. Might the presence of potentially divisive markers (caste, religion, geography, ethnicity on general) exacerbate conflictual situations? For instance, Esteban and Ray (2007a) argue noneconomic (“ethnic”) markers may play a salient role in the outbreak of conflict even when society exhibits high economic inequality and may look prima facie more ripe for a class war.

A standard tool for measuring ethnic and social divisions is that of fractionalization, roughly defined as the probability that two individuals drawn at random will come from two distinct groups. While fractionalization seems to have a negative effect on economic outcomes such as per-capita GDP (Alesina et al. (2003)), growth (Easterly and Levine (1997)), or governance (Mauro (1995)) its effect on civil conflict appears to be insignificant (Collier and Hoeffler (2004), Fearon and Laitin (2003)). Of course, as Horowitz (2000) and others have observed, it is is the presence of large cleavages that is potentially
conflictual, whereas fractionalization continues to increase with diversity. The solution is to drop fractionalization altogether. Montalvo and Reynal-Querol (2005) adapt Esteban and Ray’s (1994) measure of polarization to show that measures of ethnic and religious polarization do indeed have a significant impact on conflict (see also Do and Iyer (2006)). Obviously, more research is called for on questions such as these. For instance, it is unclear how polarization should enter an empirical specification: Esteban and Ray (2007b) argue that highly polarized societies may actually avoid conflict via deterrence, though conditional on the outbreak of conflict, polarization must vary positively with it.

The continuing study of conflict in development demands our highest priority. Certainly, the social waste of conflict dominates the inefficiency of misallocated resources which so many mainstream economists prefer to emphasize. Indeed, it is entirely possible that the much-maligned (and much-studied) inefficiencies of incomplete information are also of a lower order of magnitude. But most of all, it is the chain of cumulative causation that must ultimately drive our interest: from underdevelopment to conflict and back again to continuing underdevelopment. Conflict is one channel through which history matters.

4.4 Legal Matters

Contract enforcement, property rights, and expropriation risks: these are a few instances of legal matters that are central to development. They bear closely on that much-used catchall phrase, “institutional effects on development”. For instance, Acemoglu, Johnson and Robinson (2001) as well as the recent survey by Pande and Udry (2007) clearly have the security of property rights high on the list when discussing “institutions”. La Porta, Lopez-de-Silanes, Shleifer and Vishny (1997, 1998, 2002) and Djankov, La Porta, Lopez-de-Silanes and Shleifer (2003) begin with the premise that common (English commercial) law and civil (French commercial) law afford different degrees of protection and support to investors, creditors and litigants, and argue that it has had dramatic effects on a variety of indicators across countries: corruption, stock-market participation, corporate valuation, government interventionism, judicial efficiency — and presumably, via these, to economic indicators.

It is little surprise that the security of property rights is generally conducive to investment, and that long-term investment is especially encouraged by such security; see, e.g., Demsetz (1967). (Short-term effort on land, in contrast, may well be enhanced by insecurity of tenure.) Depending on the exact form that such rights assume, there may be further positive effects — e.g., via access to credit — that arise from the ability to mortgage or sell property (Feder, Onchan, Chalamwong and Hongladarom (1988)). Just as in the case of cross-country regressions one is invariably assailed here by standard questions of endogeneity and omitted variables. For instance, long-gestation investments may provoke — and permit — the establishment of property rights, and a high-ability agent
might use her ability to both invest and secure her rights. Nevertheless, the evidence on property rights is that by and large, they are good for investment and production (Besley (1995), Banerjee et al. (2002), Do and Iyer (2003), Goldstein and Udry (2005)), and even more obviously, property values where these are reasonably well-defined (Alston, Libecap and Schnedider (1996), Lanjouw and Levy (2002)). Instances in which property titling creates better access to credit are, intriguingly enough, somewhat harder to come by (Field and Torero (2004) and Dower and Potamites (2006) are two of the rarer examples that do document better access, but with some qualifications).

The very fact that economists have a field day establishing the effect of property rights suggests that there is a plethora of situations in which the absence of well-defined rights is the rule rather than the exception. In rural societies the world over, land rights can be highly ambiguous, and land titles can be missing even when an unambiguous definition of property exists. If one adds to this the sizeable proportion of land under tenancy, the effective security for the cultivator becomes more tenuous still (and indeed this complicates matters, because her rights may be inversely related to those of the owner!). In nonrural settings, there are substantial uncertainties for those who operate in the informal sector (such as the periodic “cleansing” of informal retailers from city sidewalks). If the above studies are to be taken seriously, there are substantial production losses from such states of insecurity.

If imperfections of the law are so inimical to the fortunes of cultivators and producers (and especially for the small and the poor among them), why do we see such institutional “failures” in equilibrium? The Coase-Posner view would presumably have none of this: in their view, legal systems would invariably develop to maximize social surplus. But of course, there could be several reasons for the persistence of “inefficient institutions”. When sidepayments are not feasible or credible, economic agents often prefer a larger share of a reduced pie to a smaller share of a more efficient pie. For instance, domestic businesses which can rely on a trusted network of kin or extended family might prefer an ambiguous legal system, so that it prevents entry. Or workers might prefer imperfect enforceability of a work norm, so that efficiency wages need to be paid. Borrowers might prefer that loan repayment cannot be fully enforced, so that incentives to repay must be built into the loan contract. And when tenancy is widespread in agriculture, the very design of overall property rights to maximize efficiency can be a highly complex problem.

The last three examples possess another feature that is worth some emphasis. It is that ambiguous property rights often have equity effects that don’t go the same way as efficiency-minded economists would like them to go (Weitzman (1974), Cohen and Weitzman (1975), Baland and Platteau (1996))). The ambiguity of property rights can serve as insurance, buffer, or redistributive device. As examples, consider broad access to water resources or grazing land, or the efficiency-wage premia that may need to be paid to workers or borrowers.
Most important, the ambiguity of property rights slows down the emergence of an overt assetless class, which has its own social payoffs (it should not be forgotten that the flip side of unambiguous rights is exclusion). For example, Goldstein and Udry (2005) develop this point of view in the context of rural Ghana, arguing that the ambiguity in property rights prevented the outbreak of extreme poverty (and had an interesting efficiency effect in the bargain, as individuals were reluctant to leave the land fallow — an important investment — in the fear that this would signal a lack of need for land).

The political economy of rights is a messy business, but of central importance in development economics. Poverty in general enhances the social and political need for ambiguity, while to the extent that such ambiguity wears on efficiency, we have an extremely important instance of nonconvergence. Sometimes such nonconvergence assumes particularly dramatic form. In West Bengal (India) “Operation Barga” provided widespread — and welcome — use-rights to registered sharecroppers (see, e.g., Banerjee, Gertler and Ghatak (2001)). Those very use-rights now lie at the heart of recent difficulties in converting agricultural land in India for use in industry. In the world of the second-best, few policies have unambiguously one-directional effects.

5 A Concluding Note: Theory and Empirics

While I have tried to provide a conceptual overview in this essay, recent research in development economics has been almost entirely empirical. A veritable explosion in computing power, the expansion of institutional datasets and their increased availability in electronic form, and the growing ease of collecting one’s own data has bred a new generation of development economists. Their empirical sensibilities are of a high order; they are extremely sensitive to issues of endogeneity, omitted variables, measurement error and biases induced by selection. They are constantly on the search for good instruments or natural experiments, and when these are hard to find, they are adept at creating experiments of their own.

There is little doubt that we know little enough about the world we live in that it is often worth finding out the simple things, rather than continue to engage in what some would term flights of theoretical fantasy. Are people really credit-rationed? Does rising income automatically make for better nutrition and health? If we had the option to throw in more textbooks, or reduce class size, or add more teachers, or install monitoring devices to track teacher attendance, which one should we do? Do women leaders behave differently from men in the policies that they adopt? Do households behave as one frictionless unit? Or, if one is the big-picture sort, have countries indeed converged over the last 200, 500 years? Are richer countries more democratic? How many excess female deaths have occurred in China or India because of gender bias? Are poorer countries more
“corrupt”?, and so on. The list is practically endless.

Why can’t well-trained statisticians answer these questions?, the somewhat churlish theoretically-minded economist might ask. Why do we need economists, who are sup posed, at the very least, to combine two observations to form a deduction? The answer, at one level, is very simple and not overly supportive of the churlish theorist’s complaint. While the questions are straightforward, the answers are often extremely difficult to tease out from the data, and you need a well-trained economist, not a statistician, to understand the difficulty and eliminate it. Because of the aforementioned econometric issues, not a single one of the questions asked above admit a straightforward answer. Development economists spend a lot of time thinking of inventive ways to get around these problems, and it is no small feat of creativity, dedication and extremely hard work to pull off a convincing solution.

It is true that the very desire to obtain a clean, unarguable answer — with its attendant desire to have control over the empirical environment — sometimes narrows the scope of the inquiry. There is often great reluctance to rely on theoretical structure (for such reliance would contaminate the near-lexicographic desire for an unambiguous result). This means that the question to be asked is often akin to that for a simple production function (e.g., “do students do better in exams if they are given more textbooks?”) or is focused on the direct effect of some policy intervention (“does the provision of health checkups improve health outcomes?”) So it is that a boring but well-identified empirical question will often be treated with a great deal more veneration (especially if a clever instrument or randomization device is involved) than a model which relies on intuitive but undocumented assumptions.

That said, it is also a fact that we know very little about the answers to some of the most basic questions, such as the ones we’ve listed above. The great contribution of empirical development microeconomics is that we are building up this knowledge, piece by piece. Whether the search for that knowledge is informed by theory or not, there will be enough theorists to attempt to put these observations together. There will be enough empirical researchers to keep generating the hard knowledge. Development economics is alive and well.

Bibliography


Coate, S. and M. Ravallion. 1993. Reciprocity Without Commitment: Characteriza-


Dube, O. and J. Vargas. 2006. Are All Resources Cursed? Coffee, Oil and Armed Conflict in Colombia. Documentos de CERAC 002748.


