DEVELOPMENT ECONOMICS THROUGH THE LENS OF PSYCHOLOGY

Sendhil Mullainathan[1]

(MIT and NBER)

[1] I am grateful to George Loewenstein and an anonymous referee for comments. Parts of this paper are also clearly inspired by discussions about current work on related topics with Marianne Bertrand and Eldar Shafir, Abhijt Banerjee and Richard Thaler.

ABSTRACT

Economists conceptualize a world populated by calculating, unemotional maximizers. This view shapes our understanding of many crucial elements of development economics, from how rural villagers save to how parents decide on whether to send their children.

Psychological research, however, has documented the incompleteness of this perspective. Individuals have self-control and time inconsistency problems. They can give into short-run temptations and later regret it. They can have strong feelings about others that drive them to commit both generous acts and spiteful ones. They often passively accept defaults, rather than making active choices. They let the institutions around them make their choices for them. They may misread new data in a way that fits what they believe in. In short, the rational maximization model may not be a very good approximation to human behavior.

In this paper, I present some of this psychological evidence that I think helps to better understand a few core issues in development economics, such as savings, education and property rights. This evidence gives us new interpretations for a variety of behavior in these contexts. It also enriches the set of policy tools we should consider. Not only does it suggest some dramatically new tools, it also suggests small cost changes to existing policies that may dramatically improve their efficacy.

Introduction

Economists often study scarcity. Yet their conception of decision-making assumes an abundance of psychological resources. In the standard economic model, people are unbounded in their ability to think through problems. Regardless of complexity, they can costlessly figure out the optimal choice. They are unbounded in their self-control. They costless implement and follow through on whatever plans they set out for themselves. Whether they want to save a certain amount of money each year or finish a paper on time, they face no internal barriers in accomplishing these goals. They are unbounded in their attention. They think thorough every single problem that comes at them and make a deliberative decision about every one. In this and many other ways, the economic model of human behavior ignores the bounds on choices (Mullainathan and Thaler 2001). Every decision is thoroughly contemplated, perfectly calculated and easily executed.

A growing body of research interprets economic phenomena with a more modest view of human behavior. In this alternative conception, individuals are bounded in all of these dimensions (and more). Practically, this conception begins with the rich understanding of human behavior that experimental psychologists have developed through numerous lab and some field experiments. This view, ironically enough, emphasizes the richness of behavior that arises from scarcities, emphasizing the bounds on cognitive and computation ability, self-control, attention and self-interest. Theoretical models are now being constructed that help to incorporate these ideas into economic applications. Even more compelling, perhaps, is the recent empirical work that suggests the importance of these psychological insights for real behavior in contexts that economists care about. In a variety of areas, from asset pricing to savings behavior to legal decision making, well-crafted empirical studies are challenging the traditional view of decision making.

This paper attempts to give a flavor of this research to those interested in development economics. I have chosen psychological insights that I think are helpful in understanding a few phenomena in development economics: parents' schooling decisions, savings, financial institutions, bureaucratic corruption, and property rights. For each of these I describe a small piece of the psychology that may be potentially relevant. In this way, I hope the reader gets a flavor of the psychological and associated field evidence while at the same time glimpsing the practical relevance of this evidence. Given the space considerations, my goals are modest. I am clearly not comprehensive in my review of the different areas of relevant psychology; that would take a book at the least. Nor am I comprehensive in describing the various psychological insights that may help in understanding any one topic (e.g. savings). As I stated, my goal is instead to give only a flavor of each topic.

Two important caveats are in order. First, there are many reasons to believe that these psychological factors may be unimportant in economic contexts. Some economists could argue that the experiments are "weak" because people in them are not financially motivated. Others might argue that market competition or arbitrage guarantee that these "irrational" choices should have no impact on economic outcomes. Yet others might argue that learning would remove these problems. I will not address these objections because they have been dealt with at great length elsewhere.[2] I am more pragmatic in my approach. I do not believe that any set of lab experiments alone can ever be a firm basis for policy. Even the best experimental evidence will face questions of context specificity, behavioral adaptation and equilibrium. Instead, these experiments are wonderful because they inspire different perspectives on old problems and new ideas for policy. Their ultimate success, however, depends on how they fare when tested in the field. So, for my purposes, the evidence I provide is merely to inspire (and not substitute) for careful tests in relevant contexts. The experimental evidence, therefore, need only pass a lower hurdle: is the bulk of the evidence sound enough to merit future empirical work or policy experimentation? The accumulated evidence, I feel, easily passes this hurdle.

Second, my attempts to incorporate psychology into development should not be confused with pejorative attempts to label the poor as "irrational". It is neither an attempt to blame the poor for their poverty nor to argue that the poor have specific irrationalities. Instead my goal is to understand how problems in development might be driven by general psychological principles that operate for both poor and rich alike. When I speak of self-control, for example, I am speaking of those self-control problems that exist in equal measure around the world. These problems may matter more for the poor because of the context in which they live but their core is a general one (Bertrand, Shafir and Mullainathan 2003).

Immediate Barriers to Education

The rational choice model of schooling is straightforward (Becker 1993). Individuals trade off the costs and benefits of schooling to decide how much schooling to get. Benefits come in a variety of forms such as better jobs or better marriage prospects. Costs could be direct financial costs (e.g. fees) as well as any opportunity costs (e.g. foregone labor). In the case of children, of course, parents make the actual choices. They do so to maximize some combination of their own and their children's long run welfare, with their exact weight depending on their altruism.

This view of education abstracts from the richness of the hardships faced by a parent trying to educate their children in a developing country. Consider a poor father in a village who is eager to send his son to school for the next school year. He recognizes the value of his son's education in allowing to him to get a

[2] See Mullainathan and Thaler (2001) for references and a summary discussion.

government job, marry better or simply exist more comfortably in a rapidly changing world. To ensure that he has money for school fees, textbooks or perhaps a school uniform, he begins to save early. Quickly he encounters some competing demands on the money. His mother falls ill and needs money to buy some anesthetics to ease her pain. Though his mother insists that her grandson's education is more, he is torn. Enormous willpower is required to let his mother suffer while he saves money that he knows could ease her pain. Knowing that he is doing what is best in the long run is small consolation in the moment. He overcomes this hurdle and enrolls his son. After some weeks, his son starts to show disinterest. Like most children everywhere, sitting in a classroom (and not an appealing one at that) does is not very appealing, especially since some of his friends are outside playing. In the evenings, exhausted from tiring physical work and feeling the stresses of everyday life, how will he handle this extra stress? Will he have the mental energy to convince his son of the value of education? Will he have the energy to follow up with the teacher or other students to see whether he as actually been attending school? This fictional example merely illustrates one important tension. Even the best of intentions may be very hard to implement in practice, especially in the high stress settings that the poor inhabit.

These problems are intimately related to how people view tradeoffs over time, a topic that psychologists and behavioral economists have studied extensively through experiments. I now describe a variety of such evidence and then return to how this evidence may help us to understand the schooling decision.

Would you rather receive $15 today or $16 in 1 month? More generally, how much money would I need to give you in one month to make you indifferent to receiving $15 today? What about in 1 year? What about in 10 years? Thaler (1981) presented these questions to subjects and found median answers of $20, $50 and $100. While at first blush these answers may seem somewhat reasonable, they actually imply huge discount rates: 345% over one month, 120% over a 1-year horizon and 19% over a 10-year horizon.[3] Subjects greatly prefer the present to the future.

These choices also imply that the rate of time preferences *changes* with the horizon. This is made most clear in the following choice problem:

Would you prefer $100 today or $110 tomorrow?

Would you prefer $100 thirty days from now or $110 thirty-one days from now?

[3] One reason subjects may be show such preferences is that they may doubt that they will actually get the money in the future, leading them to value it at a lower rate. While this may be an effect, the literature on discounting finds similar results even when these issues of trust are dealt with (Frederick, Loewenstein and O'Donoghue (2002).

Many subjects give different answers to these two questions. To questions such as the first one they often prefer the immediate reward ($100 today). To questions such as the second one they often prefer the delayed reward ($110 in thirty-one days).

Such preferences are inconsistent with the standard model. To see this, suppose people discount the future at a rate δ. Then the value of $100 today is *u(100)* and the value of it tomorrow is *u(110)*. On the other hand, in problem two the value is $\delta^{30}u(100)$ versus $\delta^{31}u(110)$. This is the exact same trade-off. In other words, with the standard constant discounting individuals should choose the same thing in both situations.

Differences in preferences for the immediate versus the future can also be seen in the field. Read, Loewenstein and Kalyanaraman (1999) ask subjects to pick three rental movies. The subjects either pick one by one for immediate consumption. Or they pick all at once for the future. When picking sequentially for immediate consumption, they tend to pick "low-brow" movies. When picking simultaneously for future consumption, they tend to pick "high-brow" movies. Once again, when planning for the future I am more willing to make choices that have long-run benefits (presumably "high-brow" movies) than when choosing in the present.

The difference in choices at different horizons poses a problem for the individual. Consider a concrete example. Suppose my preference is that next Monday I begin writing a paper rather than put that off until Tuesday. Of course, today I am busy and would rather put the paper off. What happens on Monday? What was a decision about the distant future (where I exhibited patience) becomes a decision about the present (where I exhibit impatience). My choice may now change. Once again, the option of putting it off for a day seems appealing, as appealing as it did last week when I made the same decision. In other words, there is a conflict between what I plan to do in the future and what I would actually do when the future arrives.

This conflict is one (though surely not all) of the difficulties parents face in getting their children educated. In the example I gave the father wanted his son educated and was willing in the future to put in the effort and money needed to see that happen. Yet in the moment, many immediate pressures impinge on his time, money and energy, making it hard for him to implement his plan. This view presumes that parents would like to see their children educated but simply can't find a credible way to stick to that plan. I think this perspective improves our understanding of many components of education.

It provides explanation of the gap between parents' stated goals and outcomes. The Probe report on basic education in India finds that many parents are actually quite interested in education. Even in the poorest, worst education states in India, they find over 85% of the parents agreeing that it is important for children to be educated. In the same survey, 57% of parents feel that their sons should study "as far as possible". Another 39% feel that they should get at least a grade 10 or grade 12 education. Clearly parents in these areas value education. Yet this contrasts with very low educational attainment in these states. This gap is reminiscent of the gap between desired and actual retirement savings in the United States. In one survey, 76 percent of Americans believed that they should be saving more for retirement. In fact, 55% felt they were behind in their savings and only 6% reported being ahead (Farkas and Johnson 1997). Though they want to save, many never make it happen. As noted earlier, immediate pressures are even more powerful in the education context. Putting aside money to pay for schooling requires making costly immediate sacrifices. Fighting with children who are reluctant to go to school can be especially draining when there are so many other pressures. Walking a young child every day to a distant school requires a constant input of effort in the face of so many pressing tasks. Put another way, if a middle-class American, supported by so many institutions, cannot save as much as they want, how can we expect a Rajasthani parent to consistently and stoically make all the costly immediate sacrifices to implement their goal of educating their children?

It also helps to explain in part an interesting phenomenon in many developing countries: sporadic attendance. In contrast to a simple human capital model, education does not appear to follow a fixed stopping rule, with students going to school consistently until a particular grade. Instead, students go to school for some stretch of time, drop out for some stretch and then begin again. This sporadic attendance, though potentially far from optimal, is a characteristic of the dynamically inconsistent preferences described earlier. When faced with particularly hard to resist immediate pressures, individuals will succumb to them. When these pressures ease, it becomes easier to implement the original plan of sending their child to school and they may revert to it. Relatedly, in many discussions of self-control, the importance of salience is often emphasized (Akerlof 1991). To this end, parents who have "slipped off the wagon" may find some salient moments that encourage them to try again to get their children to school. One empirical prediction here is that at the beginning of the school year, attendance should perhaps be higher than at any other time as many parents decide to give it another try. As they succumb to immediate pressures, attendance would then decline throughout the year.[4]

This perspective also has some insights for policy. First, policies that spread immediate pressures over time could be beneficial. For example, school fees that require continuous small payments rather than one large payment may make it easier for parents to finance savings. It is far harder to have the will power to save up

[4] This last point provides one way to distinguish this explanation from a rational model with very large liquidity shocks. Moreover, in such a rational model faces difficulties if parents rationally forecast such shocks and there are scale economies to attending for long continuous periods. In this case, parents should build a "buffer stock" early on to insure against such shocks and then send the child to school for one long (and presumably more productive) stretch.

for a big purchase (such as uniforms) than to pay small fees each week or month.[5] Second, this perspective should alter policies that attempt to increase parental demand for education. For example, the success of bonus payments to parents for children's enrollment depends crucially on their structure. If the payments are paid at the end of the year, then they are unlikely to work particularly well. In this model, parents already recognize a long-run reward to education. Adding to that will do little to solve the core problem. In contrast, bonus payments that appear at much higher frequencies may help to tilt the tradeoff in the short-run, which is the real barrier. Third, programs that make schooling more attractive to students may provide a low-cost way to make it easier for parents to send children to school. For example, a school meals program may make school attendance attractive to children and ease the pressure on parents to constantly encourage their children to get to school (see Vermeesch 2003 for a discussion of such programs). One could even be creative in coming up with these programs. For example, school sports, candy or any number of other cheap inputs that make schooling more attractive to children may have large effects. In fact, under this model such programs could have extremely large benefit to cost ratios, much larger than could be justified by the monetary subsidy alone.

In my opinion, this perspective on schooling matches the complexity of life in developing countries. Of course, immediate pressures are not the only problem. Numerous other factors—from liquidity constraints to teacher attendance—surely play a role. Yet, those have been explored and are very much on the radar screen of many development economists. This force, while potentially powerful, is not and deserves more scrutiny.

Demand for Commitment and Savings

The difficulty of sticking with a course of action in the presence of immediate pressures also has implications for how individuals save. In the standard economic model of savings, there is no room for such pressures. Instead, people calculate how much money is worth to them in the future, taking into account any difficulties they may have in borrowing and any shocks they may suffer. Based on these calculations, they make a contingent plan of how much to spend in each possible state. They then, as discussed before, implement this plan with no difficulty. As noted earlier, for the poor in many developing countries, implementing such plans is much easier said than done. They face a variety of temptations that might derail their consumption goals.

Behavioral economists have recently begun to better understand the devices people may use to deal with such temptations. The inter-temporal preferences noted earlier (short run impatience, long run patience) are often modeled as a discount rates that vary with horizon. People have very a high discount rate for short horizons (decisions about now versus the future) but a very low one for distant horizons. This is often called

[5] Note that in this framework, unlike a liquidity constraint framework, this policy would work even if these payments all had to be made *prior* to the school year starting. This would be analogous to the use of lay-away in the United States.

hyperbolic discounting because the original curve used to produce it was hyperbolic in shape (Strotz 1956, Ainslie1992 Laibson 1997).

A key question in this model is whether people are sophisticated or naive in how they deal with their temporal inconsistency. Sophisticated people would recognize the inconsistency and (recursively) form dynamically consistent plans. In other words, they would only make plans that they would follow through on. Naïve people on their hand would not recognize the problem and would make plans assuming that they'll stick with them, only to abandon them when the time comes. There are reasons to believe both views. On the one hand, individuals appear to consciously demand commitment devices, to help them commit to a particular path. On the other, they appear to have unrealistic plans. Perhaps the best fit of the evidence is that individuals partly (though not necessarily fully) recognize their time inconsistency.

The practically important feature of this view is that the commitment implicit in institutions is very important for understanding behavior. Institutions can help solve self-control problems by committing people to a particular path of behavior. A common analogy here is with Ulysses, who ties himself to his ship's mast so that he can listen to the song of the sirens but no be lured out to sea by them. While not so dramatic, similar commitment devices exist in everyday life. Many refer to their gym membership as a commitment device ("Being forced to pay that much money every month really gets me to go to the gym lest I waste it".). Or to take another example, christmas clubs, while less common than before, used to be a very powerful commitment tool for saving to buy Christmas gifts.

Suggestive evidence on the power of commitment devices is given in Gruber and Mullainathan (2002) which studies smoking behavior. Rational choice models of smoking treat it roughly like any other good. Smokers make rational choices about their smoking understanding the physiology of addiction that nicotine entails. Behavioral models on the other hand recognize a self-control problem in the decision to start smoking and in the decision (or rather attempts) to quit. There is some survey evidence suggestive of the behavioral model. Smokers often report that they would like to quit smoking but are unable to do so. This resembles the temporal pattern above. Looking into the future, smokers would choose to not smoke. But when the future arrives, they end up being unable to resist the lure of a cigarette today (perhaps promising that tomorrow they quit). To differentiate these theories we examined the impact of cigarette taxes. Under the rational model, smokers are made worse off. This is a standard dead-weight loss argument. Smokers who would like to smoke cannot now because of the higher price. In models with time hyperbolic discounters, however, taxes could smokers made better off. The very same force—high prices driving smokers to quit—that is bad in the rational model is good in this model. Because smokers wanted to quit, but were unable to, they are now better off. In the parlance of time-inconsistency models, the taxes serve as a commitment device.

To assess well-being we use self-reported happiness data. While such data are far from perfect, they can be useful especially in contexts such as these where the variable of interest is relatively clean and thereby the mis-measurement is simply absorbed in the residual. Using a panel of states in the United States, we find that happiness of those who tend to smoke increases when cigarette taxes increase. Relative to the equivalent people in other states (and relative to those who tend not to smoke in their own state), these people show actual rises in self-reported well-being. In other words, contrary to the rational model and supportive of the behavioral model, cigarette taxes actually make those prone to smoke *better off*. This kind of effect is exactly the one I alluded to in the introduction: institutions (cigarette taxes in this case) have the potential to help solve problems within people as well as between people.

There is also evidence on people actively choosing commitment devices. Wertencroch (1998) argues that people forego quantity discounts on goods they'd be tempted to consume (e.g. cookies) in order to avoid temptation. This is a quantification of the often-repeated advice to dieters: don't have big bags of cookies at home. If you must buy tempting foods, buy little of it. Trope and Fischbach (2000) show how people strategically use penalties to spur unwanted actions. They examine people scheduled to undertake small, unpleasant medical procedures and show how these people choose to voluntarily take on penalties for not undertaking the procedure. In fact, they smartly pick these penalties, by picking higher ones for more aversive procedures. Ariely and Wertenbroch (2002) provide even more direct evidence. They examine whether people use deadlines as a self-control device and whether such deadlines actually work. Students in a class at MIT chose their own deadlines for when to submit three different papers. The deadlines were binding so in the absence of self-control problems they should clearly choose the latest deadlines possible for all the papers. They were told there was neither benefit to an early deadline nor cost to a late one so they can only benefit from the option value of being able to turn it in later. In contrast, students chose evenly spaced deadlines for the three papers, presumably to give them incentives to get the papers done in a timely manner. Moreover, deadlines appeared to work. In a related study, they show that people given evenly spaced deadlines do better than those given one big deadline at the end.

I think savings in developing countries can also be better understood through this perspective. It provides an alternative view on institutions such a roscas (Gugerty 2003). In a rosca, a group of people meets together at regular intervals. At each meeting, members contribute a pre-specified amount. The sum of those funds (the "pot" so to speak) is then given to one of the individuals. Eventually, each person in the rosca will get their turn and thus get back their contributions. Roscas are immensely popular but what is their attraction? They often pay no interest. In fact, given the potential for default (those who received the pot early not continuing to pay in), they may effectively pay a negative interest rate. One reason for their popularity may be that they serve as a commitment device in several ways. By making savings a public act, they allow social pressure from other rosca members to commit them to their desired savings level (Ardener and Burman 1995). As some rosca participants say, "you can't save alone". Other rosca members have all the incentives to make

sure each other member continues to contribute. They also allow individuals to save up to larger amounts than they normally could achieve given their self-control problems. Imagine someone who wished to make a durables purchase (or pay school fees) of 1000 rupees. By saving alone and putting aside money each month, they face a growing temptation. When they reach 400 rupees, might not some other purchase or immediate demand appear more attractive? The rosca doesn't allow this temptation to interfere. Individuals get either nothing or the full 1000 at once. This all or nothing property might make it easier to save enough to make large purchases.

It also helps to provide a more nuanced view of individuals demand for liquidity. In the standard logic, the poor unconditionally value liquidity. After all, liquidity allows people to be able to free up cash to attend to immediate needs that arise. If a child gets sick, money is needed to pay for medicine. This might be especially true for the poor. Shocks that are small for the well off can be big for the poor and they would need to dip into real savings to address them. The poor, in these models, however face a tradeoff. They value liquidity for the reasons cited above but liquidity for them is also a curse: it allows them to too easily dip into savings. Durable goods and illiquid savings vehicles may actually be preferred to liquid savings vehicles. Cash, for example, may be far too tempting and spent too readily. On the other hand, by holding their wealth in items such as jewelry, livestock, and grain, individuals may effectively commit themselves not to give into immediate consumption pressures. In these models, therefore, there is an optimal amount of liquidity. Even when liquidity is provided at zero cost, they will choose some mix of illiquid and liquid assets.

Another implication from this perspective is that revealed preference fails as a measure of policy success. Observing that people borrow at a given rate (and pay it back) does not necessarily mean that the loan helps them. In some cases, it may because it helps them deal with a liquidity shock. In other cases, it may not because it assists them in giving way to immediate temptations, leaving them straddled with debts they must repay. This distinction is important for understanding micro-credit in developing countries. Often, the metric of success for such programs is whether they are self-sustainable. Such a metric makes sense if revealed preference makes sense. Profitability would imply that people prefer getting these loans even at a non-subsidized rate; revealed preference then implies their social efficiency. Yet in the presence of time inconsistency, profitability of micro-credit could mean very little about social efficiency. The key question is to what extent the loans exaggerate short-run impatience and to what extent they solve long-run liquidity constraints.[6] Ultimately one needs a deeper understanding of what drives borrowers. One avenue for this might be data on loan usage. Are loans being spent on long-run investments (as is often touted) or spent on short-run consumption? Of course, some short-run consumption might well be efficient but this data combined with an understanding of the institution would help to better understand (and improve) the social efficiency of micro-credit.

[6] To make this contrast stark, note that in the United States, payday loan companies are a very profitable form of micro-credit.

Policy can also provide cheaper and more efficient commitment devices. After all, even if saving in grain is an expensive way to produce a commitment device. Vermin may eat the grain and the interest rate earned on the grain could be zero or even negative. Moreover, it is important to recognize, that even if people demand such commitment devices, the free market may not do enough to provide them. The highly regulated financial markets in developing countries may lead to too little innovation on these dimensions. Monopoly power may also lead to inefficient provision of these commitment devices depending on whether a monopolistic financial institution can extract more profits by catering to the desire for commitment or to the temptations themselves. In this context, governments, NGOs and donor institutions can play a large role by promoting such commitment devices.

Ashraf, Karlan and Yin (2004). give a stunning illustration. They offer savers at a bank in the Philippines the opportunity to participate in "SEED" accounts. These accounts are like deposit accounts, except individuals cannot withdraw deposits at will. Instead, they can only withdraw the money at a prespecified date or once a prescpecified goal has been reached. This account does not pay extra interest and is illiquid. In most economic models, people should turn down this offer in favor of the regular accounts offered by that bank. Yet they find large demand for them. More than 30% of those offered the accounts take them up. They also find these accounts help these individuals to save. Six months later, those offered the accounts show substantially greater savings rates. Experiments such as these will, I feel, eventually help to deepen our understanding of savings and greatly improve development policy.

Defaults and Financial Institutions

Financial institutions do not simply help savings through their commitment value. A very important set of results in behavioral economics suggests that they affect behavior simply through the status quo they produce. Samuelson and Zeckhauser (1988) document a variety of phenomena known as the status quo bias. Here is a simple example. A group of subjects is given the following choice:

> *You are a serious reader of the financial pages but until recently have had few funds to invest. That is when you inherited a large sum of money from your great uncle. You are considering different portfolios. Your choices are:*
>
> *Invest in moderate-risk company A. Over a year's time, the stock has .5 chance of increasing 30% in value, a .2 chance of being unchanged, and a .3 chance of declining 20% in value.*

Invest in high-risk company B. Over a year's time, the stock has .4 chance of doubling in value, a .3 chance of being unchanged, and a.3 chance of declining 40% in value.

Invest in treasury bills. Over a year's time, these will yield a nearly certain return of 9%

Invest in municipal bonds. Over a year's time, they will yield a tax-free return of 6%.

A second set of subjects is given the same choice but with one small difference. They are told that they are inheriting a portfolio from their uncle in which most of the portfolio is invested in moderate-risk company A. The choice now is subtly changed. It is how much of the portfolio to *change* to the options above. Interestingly, they find a large difference between the two treatments: much more of the money is re-invested in A when it is the status quo.

This bias towards the status quo appears to run quite deep and is not just due to some superficial explanations (such as information content of the uncle's investments). Samuelson and Zeckhauser (1988) demonstrate this with a very interesting piece of evidence from the field. In the 1980s, Harvard University added several plans to its choice of health plans. This provides an interesting test of status quo bias: how many of the old faculty chose the new plans and how many of the newly joined faculty chose the plan? They find a stark difference. Existing employees "chose" the older plans at a two to four times higher rate. In other words, incumbent employees make the easiest choice of all: to do nothing.

This bias towards the status quo could perhaps be motivated towards the deeper phenomena of automatic behavior. Psychologists have recently documented numerous instances of the idea that people often make automatic, non-conscious choices. Gilbert, Tafarodi and Malone (1993) give an illustrative example of automaticity. Subjects are exposed to information about a criminal defendant that were false. On some trials subjects were exposed to these sentences while cognitively loaded with another task or while under time pressure. In these conditions subjects automatically assumed the (false) statements to be true rather than examining them. This illustrates one of the basic ideas behind this research on automaticity. Unless attention is consciously drawn to a decision, it will be made through some automatic processes. In many practical situations, the likely automatic

process is to simply do nothing. Thus, what economists view as a "choice" may really not be an active choice at all. It may instead reflect default behavior combined with the institution underlying that choice.

Madrian and Shea (2001) conducted a particularly telling study along these lines. They studied a firm that altered the choice context for employee participation in their retirement plan. When they join the firm employees are given a form that they must fill out in order to participate in the savings plan. Though the plan is quite lucrative, participation is low. Standard economic models might suggest that the subsidy ought to be raised. This firm instead changed a very simple feature of its program. Prior to the change, new employees received a form that said something to the effect of "Check this box if you would like to participate in a 401(k). Indicate how much you'd like to contribute". After the change, however, new employees received a form that said something to the effect of "Check this box if you would like to **not** to have 3% of your pay check put into a 401(k)". By standard reasoning, this change should have little effect on contribution rates. How hard is it to check off a box? In practice, Madrian and Shea (2001) find a large effect. When the default option is to not contribute, only 38% of those contributed. When the default option was contribution, 86% contributed. Moreover, even several years later those exposed to a contribution default still show much higher contribution rates.

These results are consistent with (and motivated) the discussion above. While we cannot be sure from this data what people are thinking, I would speculate that some combination of procrastination and passivity played a role. Surely many looked at this form and thought, "I'll decide this later". But later never came. Perhaps they were tempted by more other than deciding on 401(k) contribution rates (hard to believe, but there are more interesting activities). Perhaps it simply slipped out of their attention because other factors came to occupy it. In either case, whatever the default on the form was, they ended up with. In fact, as other psychology tells us, as time went on they may well have justified their "decision" to themselves by saying, "3% is what I wanted anyway" or "that 401(k) plan wasn't so attractive. In this way, their passivity made their decision for them. By making the small active choice to choose later, they ended up making a large decision about thousands of dollars of money.

These insights can also help us design whole new institutions. One example is Save More Tomorrow, a program created by Thaler and Benartzi (2003). The basic idea of Save More Tomorrow is to get people to make one active choice but to have them make it in such a way that if they remain passive afterwards they are still saving. To participate, they decide on a target savings level (and we know from before that people actually want to save). Once they decide on what they'd like to save, they agree to start deductions at a small level from their paycheck *next year*. And then each year, as they receive a raise their deductions increase until they hit their target savings level. They can opt out of the program at any time. But the cleverness of the

program is that if they do nothing and remain passive, they will continue to save (and even increase their savings rate).

The results have been stunning. In one firm, for example, more than 75% of those offered the plan participated in it rather than simply trying to save on their own. Of these, interestingly few of them (less than 20%) opted out. As a result, savings rates went up sharply. By the third pay rise (as the default increases cumulated), individuals had more than tripled their savings rates. But perhaps the greatest success has been the diffusion of this product. Many major firms and pension fund providers are thinking of adopting it and participation in the program will likely soon number in the millions. Save More Tomorrow is an excellent example of what psychologically smart institutional design might look like in the future. It does not solve a problem between people but instead helps solve a problem within the person: not saving as much as they would like.[7]

One simple implication of these results is that behavior should not be confused with dispositions (Bertrand, Shafir and Mullainathan 2004). An economist observing the middle-class American and a rural farmer's savings behaviors might be tempted to conclude something about different discount rates. The high savings of the middle-class American surely reflects their greater patience. As we have seen, this need not be the case. Such an inference could be just as wrong as inferring that those who defaulted into 401(k) are more patient than those who were not defaulted into it. The behavioral difference may be that better institutions facilitate more automatic, default savings.

Another implication is in the form of banking reform. The lessons learned in the United States could easily be transferred to parts of developing countries. First, institutions such as automatic deposit of paychecks (as well as the ability to reroute some of the money directly to a savings account) could be very powerful in spurring savings. Banking innovations such as these could be very inexpensive yet have profound effects on the savings rates of the middle-class in developing countries.

Second, it suggests that the simple extension of banking to rural areas could in and of itself have large impacts on behavior. While this is not as powerful of a default as having your paycheck automatically deposited, it may very well help to have the money out of easy access. The person then has to make one active decision—putting the money into the account—but then the act of keeping the money becomes a passive one. When money is around, active effort is required to save it away. When money is in the bank account, active effort is required to go and get it in order to spend it. In this sense, a bank account may serve as a very weak commitment device. By keeping the money at a (slight) distance, spending it may be a lot less tempting.

Loss Aversion and Property Rights

Consider the following simple experiment. Half a room of students is given mugs and the other half are given nothing (or a small cash payment roughly equivalent to the value of the mugs). The subjects are then placed in a simulated market where a mechanism determines an aggregate price at which the market clears. How many mugs should change hands? Efficiency dictates that market clearing should allocate the mugs to the 50% of the class that values it the most. Since the mugs were initially randomly assigned, roughly half of this group should have started off with mugs and half should have started off with no mugs. Consequently, trading should have resulted in exactly half the mugs changing hands.

Kahneman, Knetsch and Thaler (1990) have in fact run this experiment. Contrary to the simple prediction, however, they find a stunningly low number of transactions. Roughly 15% of the mugs trade hands. The problem is seen if we look at how students value the mugs. Those who were given the mugs put a reservation price at three *times* those who were not given the mug. Given that, it is no surprise that so few mugs change hands. Numerous follow up experiments have been run on this so-called *endowment effect* to rule out the obvious explanations: an income effect, the value of mug recipients being able to see and feel the mug, or small transaction costs of some form. In the end, the phenomenon is robust. Those given objects very quickly appear to come to value them more than those not given the object.

This phenomenon reflects in part a deeper fact about utility functions: prospect theory. In fact the original experiment was motivated by prospect theory. In prospect theory, people's utility functions are defined in large part on changes. In the traditional model of utility people would value the mug at $u(c+Mug)–u(c)$. That is their utility is defined in absolute levels of consumption and the mug adds to that. In the prospect theory approach, utility is defined by a value function that is evaluated locally and in changes. Those who receive the mug consider its loss as a function of $v(-Mug)-v(0)$. Those who do not receive the mug value its gain at $v(Mug)-v(0)$. Notice the symmetry in the original function: both those with and without the mug value it the same (on average). In the second formulation, however, nothing guarantees the symmetry. The difference in valuation between the two depends on whether $v(Mug)$ is bigger or smaller than $-v(-Mug)$. The evidence above is consistent with a variety of evidence from other contexts: losses are felt more sharply than equivalent gains. Thus $v(x) < -v(-x)$. This phenomenon, known as loss aversion, has been seen in many contexts. Perhaps the two cleanest examples are in Odean (1998) and Genesove and Mayer (2001). Odean (1998) shows that small investors in the stock market are more willing to sell stocks they have made money on than ones they have lost money on. This fact may seem quite obvious but it is inconsistent with standard utility theory (he rules out the obvious tax explanations) since gains and losses are symmetric: investors

[7] In this short space, I cannot do justice to all the psychological tools that the SMarT plan relies on. The full discussion in the original paper is well worth reading as an example of how to use psychological tools to

should merely take the trades they view as best. In fact, he finds that this strategy of holding losers and selling winners results in negative abnormal returns. Investor's unwillingness to take on losses, on the other hand, is quite consistent with loss aversion. Another example, quite familiar to many who have owned housing is given in Genesove and Mayer (2001). They find that individuals who have taken a loss on their house set far higher prices when it comes time to sell. It appears that they are more willing to gamble to break even, a phenomenon quite consistent with loss aversion.

The insight about loss aversion can also help understand why policy change is so difficult in developing countries. Consider market reforms that transfer resources from one group to another with an efficiency gain. For example, suppose privatizing a firm will result in gains for customers while resulting in losses for incumbent workers. Under this perspective, such reforms are fought so vigorously partly because the losses are felt far more sharply by the workers. One implication of loss aversion is, at the margin, to pursue strategies that preserve the rents of incumbents rather than ones that try to buy out incumbents. All other things equal, a strategy that offers a buyout for incumbent workers will be far more costly than one that grandfathers them in. The buyout requires the government to compensate them for their loss and this can be much larger than simple utility calculations can suggest. In contrast a strategy that guarantees incumbent workers a measure of job security would not need to pay this cost.[8] Many situations of institutional change require some form of redistribution. The recognition of loss aversion suggests that successful policies may require protection the losses of incumbents.

Loss aversion also reinforces the importance of well-enforced property rights. Consider a situation where there is a single good, such as a piece of land *L*. Suppose that there are two individuals A and B who can engage in force to acquire or protect the land and that engaging in violence may result in acquisition. In the presence of well-defined property rights (say this land belongs to person A), the decision to engage in force is straightforward. If B engages in force he stands to gain $v(L)$ if his force is successful. A on the other hand stands to lose $v(-L)$ if he doesn't engage in force. In this case, loss aversion implies that A stands to lose a lot more than B could gain. So with well-defined property rights A would engage in more force than B. Conseequently, B may never attempt force. So even in the absence of enforcement, loss aversion may mean that well-defined property rights may deter violence.

Consider now the case of ill-defined property rights. Suppose that both are unsure who owns the land. Specifically take the case where they both think they own it. This is an approximation to the situation where ownership with probability ½ already gives a partial endowment effect, or to the situation below of biased beliefs where both parties may have probability greater than ½ of owning it. In this case, both A and B think they stand to lose $v(-L)$ if they do not fight for the land. In other words, in the absence of well-defined

---

better design policy.

property rights, both parties will put in large amounts of resources to secure what they already think is theirs. This to me is one of the powerful implications of loss aversion. Appropriately defining property rights prevents two (or more) parties from having an endowment effect on the same object. Conflicting endowments such as this are sure to produce costly attempts at protecting the perceived endowments. Anything from costly territorial activities (fencing and de-fencing) all the way to violence may result.

Social Preferences and Teacher Motivation

In many important development contexts, self-interested behavior is very deleterious. Bureaucrats in many countries are corrupt. They enforce regulations sporadically or take bribes. Another stark example is teacher absenteeism. Numerous studies have found that teacher absenteeism is one of the primary problems of education in developing countries. Teachers simply do not show up to school and as a result little education can take place. This blatantly selfish behavior stands in contrast to some evidence on social preferences that individuals may value the utility of others. I will review this literature and describe how social preferences may be contributing to the problem and may serve as part of the solution.

A very simple game called the ultimatum game has become an excellent tool for studying social preferences (Guth, Schmittberger and Schwarze 1982, Thaler 1988). In this game, one player (call him the Proposer) makes the first move and offers a split of a certain amount, say $10. The second player ("Responder") decides whether to accept or reject this split. If it is accepted, P and R get the proposed split. If rejected, then both get zero. What makes this game so intriguing is that it clarifies two interesting issues in interpersonal preferences. First, will the responder accept "unfair" offers? In the pure self-interest model the responder should accept any offer greater than 0 and be indifferent with even a zero offer. Second, what kind of offer will the proposer make given the responder's rejection strategy? Is the proposer motivated only by the threat of rejection? Of course in the pure self-interest model, he would offer a tiny bit above zero (or even zero itself) knowing that there's no fear of rejection.

[8] Of course, this is a comparative static only. In any given context, there may be pressing reasons to favor one policy over the other.

This game has been run in many, many countries and for stakes that range from a few dollars in the US to a few months of income in many countries. Yet the pattern of findings is relatively constant.[9] First, responders often reject unfair offers (i.e. away from 50-50 splits). Second, proposers often make very fair offers, close to 50-50 or 60-40. Moreover, proposers' fair offers are not just driven by fear of rejection. They tend to make offers larger than a simple (risk-neutral) fear of rejection implies. This is most directly seen in a variant of the Ultimatum game, called the dictator game. Here the Proposer makes an "offer" but the Responder has no choice but to accept it. In this game, the threat of rejection is removed and one continues to find non-zero offers by the Proposer, though they are lower than in the Ultimatum game.

The ultimatum game illustrates two facts about interpersonal preferences. First it and the Dictator game suggests (rather prosaically) that people care about others. These are one-shot games with no chance for repetition. Yet people give away rents to others. Such "altruistic" preferences are used to a limited extent in economics (often within a family or perhaps village). Yet here we see them pretty universal. This, of course, to most people is not much of a surprise. The large amounts of charitable giving that occurs in most societies, the volunteer activity and the spending of private time on public goods (e.g. recycling) all point at such preferences.

Reciprocity often underpins such preferences as illustrated in a very nice experiment by Regan (1971). Subjects in this study rate the quality of some painting along with another person (who is actually a confederate). Partway through the experiment, during the rest period, the confederate leaves the room. When he comes back he has a Coke for himself and also has brought one for the subject. In a control condition, he merely leaves the room and comes back (with no Coke for himself or the subject). So subjects receive an unsolicited act of kindness while others do not. At the end of the experiment, as they are parting ways, the confederate mentions to the subject that he's selling raffle tickets and that he wins a prize if sells more tickets than anyone else. Could you help me and buy some tickets he asks the subject. This is the outcome of interest in this experiment: how many tickets does the subject buy. Relative to the control condition, he buys far more tickets if the confederate did the small, unsolicited, favor of buying the subject a Coke. In fact, so big is the effect that the return on the favor is quite large. He bought a 10-cent can of coke and ended up buying at least two more raffle tickets at 25 cents. Consequently for a 10-cent "investment" he yielded 50 cents.[10] Such reciprocal fairness is ubiquitous. Survey firms use it by paying people *prior* to filling out the survey because they realize that the norm of reciprocity binds individuals to send the form back. Nonprofits use it by sending small "gifts" along with their request for donations. The reciprocity norm is one specific and ubiquitous form of altruistic preferences.

[9]See Heinrich et. al., (2002), however, for some interesting differences in some tribal cultures.

Another very important wrinkle to the altruism perspective is provided by experiments in helping behavior. Darley and Latane (1968) for example did a study at Columbia where subjects believed they were in a roundtable, virtual conversation. They were sitting in a room with a mike and speakers. They were told that the conversation was with either one other person or with six other people and were told that the conversation would go in turns with only one person's mike functioning at any given time. Partway through the "conversation", the subject hears the speaker go through a seizure of some form and make a request to get help from the experimenter. When the person feels they are the only other listener, most (though surprisingly not all) seek help. When the person feels there are other listeners, nearly none seek help. Experiments such as this underline the potential fragility of pro-social behavior. It is by no means universal and is importantly shaped by context.

Yet the second fact, rejection by the Responder, points at an equally important fact about inter-personal preference. People will pay costs themselves in order to punish those they feel are being unfair.[11] By rejecting an offer, the Responder is passing up money to punish the Proposer. This type of behavior has both illustrates part of the "dark side" of interpersonal preferences. In simple altruistic models, interpersonal preferences are only a good thing: having one person care in a positive way about another only makes it easier to deal with externalities and so on. The Responder's behavior shows, however, the way in which could potentially cause inefficiencies and conflicts.

This possibility is clearest in a classic experiment by Messick and Sentis (1979). They ask subjects to imagine they have completed a job with a partner. They are asked to decide what "fair" pay for their work is. They divide the subjects into two groups however. One group is told to imagine that they had worked 7 hours on the task while the partner had worked 10. The other group is told to imagine that they had worked 10 hours while the partner had worked 7. Both groups were told that the person who had worked 7 hours had been paid $25 and were asked what the 10 hour person should be paid. Those who were told that they had worked 7 hours (and paid $25) tended to feel that the 10-hour subject should be paid $30.29. Those who were told that they had worked 10 hours, however, felt they should be paid $35.24. The source of the bias can be seen in the bimodality of the distribution of perceived "fair" wages. One mode was at equal pay ($25 for both) while the other mode was at equal hourly wage (so the 10 hour worker gets paid approximately $35.70). Interestingly, the difference between the two treatments was mainly in the proportion in each mode.

[10] Of course, the effect may would have been smaller had subjects perceived Joe as having bought the Coke for purposes of an investment.

[11] One of the debates in the experimental literature in economics is whether this "punishment" view is needed to explain these data. There is enough auxiliary evidence, however, that while the punishment view may not be the full story it is at least part of the story.

Those who had worked 7 hours showed more subjects at the equal pay level mode while those who had been told they'd worked 10 hours shoed more subjects at the equal hourly pay mode. In other words, both groups recognized two compelling norms: equal pay for equal work and equal pay for equal output. Yet their roles determined (in part) which norm they picked.

These results extend beyond picking between two fairness norms. Such conflicts could easily arise even if there's disagreement about measuring input levels (which often are not fully observed). They speak to the source of a problem created by fairness. When there is not universal agreement about the fair division, "fairness" preferences can very quickly create conflict.

These experiments as a whole illustrate the complexity of social preferences. Individuals in some contexts do much to help others (at great costs to themselves). Reciprocity especially appears to be a powerful force. They will also at cost to themselves punish those they think are being "unfair". The final behavior is especially important since notions of fairness are often driven by self-interest.

Let us return to the case of teacher absenteeism. The PROBE report surveyed teachers extensively in many areas of India and noted high absenteeism levels. It's in-depth and interviews are illuminating about their attitudes. For example it notes:

Having said this, the main issue may not be the low initial motivation of teachers as the fact that many of them lose their motivation over time. Indeed, among recently appointed teachers we often met people with genuine enthusiasm. The honeymoon, however, is usually short-lived, as the morale of young teachers is battered day after day.

Much of this psychological battering can be viewed as a perceived failure of reciprocity. As noted earlier, individuals strongly adhere to the norm of reciprocity. Failures of reciprocity (or perceived failures) can result in punitive or self-interested behavior in response. Teachers may feel a strong social preference early on and be motivated to teach and give much more than they need to. After all, from a pure self-interest motive, they know they can get away with very little teaching. Yet they may be initially motivated to do more, to come to school, to struggle with tougher students and so on. They may view these contributions as a "gift". One reason for this of course is the initial framing of the job (as a "plum job, with good salaries, secure employment and plenty of other time for other activities"). Thus, a young teacher may think, "I am giving a lot to the school". As with any giving, however, the teacher may expect strong reciprocity and see (perhaps in a self-interested way), many outcomes as a lack of reciprocity. For example, the Probe report notes that:

The most common complaint is that schools are under-equipped, under-funded, under-staffed and over-crowded. Poor infrastructural facilities were mentioned by 63 percent of teachers as one of the problems they face.

So teachers may feel that the government is not reciprocating their "gifts". This may be especially exaggerated by the transfer system in India:

Unwanted postings and arbitrary transfers are seen as a constant threat. Teachers spend a great deal of time and energy trying to avoid undesirable transfers, lobbying for preferred postings, and building up influential connections to play the transfer game.

Thus both the benign neglect of schooling and the active transfers could easily drive teachers to feel that the government does not reciprocate their efforts. They may also come to feel similarly vis a vis parents:

Teachers are often frustrated by the apathy of parents towards their children's education. They complain that parents do not send their children to school regularly, or withdraw them for flimsy reasons. They also see much foot-dragging even when children are at school: parents send them late and in tattered clothes, try to dodge the fees, and generally fail to watch their children's needs and progress. As teacher perceive it, their own efforts to keep the children at school are not reciprocated by the parents.

Thus, even an initially motivated teacher may very quickly feel justified in their apathy. They gave it their best and think that their efforts were not reciprocated. Are these inferences justified? Perhaps not. As in the Messick and Sentis study teachers may very well be making such inferences in a self-interested way. The failure of the context may very well be in it allowing teachers to make such biased attributions of fairness. Alternatively, teachers may very well be justified in these attributions. We simply cannot tell.

In either case, this perspective suggests that the problem of teacher attendance cannot be studied in isolation. Policies that affect school resources or student attendance may have a large, indirect effect on teacher attendance. More realistically, the impact of teacher incentive policies may vary dramatically with the context. In a context of limited resources where attendance is low, these policies may have only a small or moderate impact. On the other hand, if they are coupled with other policies to increase resources as a whole and student attendance, their impact might be much larger. The teachers would then no longer self-justified for their absence and the incentives needed to get them to work may be much smaller.[12] Of course, I suspect that it is the new teachers were the effects might be biggest. Amongst existing teachers, it is harder to tell

[12] Part of this implication might be counter-intuitive from a pure self-interest view. For example, it may be easier to get teachers to come to school if attendance is high than when it is low. This would appear paradoxical if teachers were simply trying to reduce the amount of work they were doing since higher attendance would precipitate even more work when they show up.

whether they will anchor on past non-reciprocity or adapt to the new context. While other factors clearly play a role in driving teacher absenteeism, a deeper understanding of their social preferences will, I think, also help to solve the problem.

Norms and Inequality

In 1937, Sherif brought subjects into an interesting psychophysics test. They were seated in a totally dark room facing a pinpoint of light some distance from them. After some time where nothing happens, the light appears to "move" and then disappear. Shortly thereafter, a new point of light appears. It too moves after some time and then disappears. Interestingly, this movement of the light is a pure psychophysical phenomenon known as the autokinetic effect. The light does not actually move, the eye merely makes it appear to move. The subjects were put in this context for repeated trials (many different resets of the light) and asked to estimate how far it had "moved". When performed alone, these estimates were variable, ranging from an inch to several feet. However, an interesting pattern developed when subjects did this task in groups of two or three. Subjects' estimates invariably began to converge on a particular number. A group norm quickly developed. In one variant, a member of the group was a confederate, someone who worked for the experimenter and gave a specific number. He found that the subject quickly converged to the confederate's answers. Others have found that norms manipulated in this way persist for quite some time. Even when subjects are brought in up o a year later, they show adherence to that initial norm. Moreover, within the context of the experiment, Jacobs and Campbell (1961) has shown how norms can be transmitted across "generations" of subjects. Suppose subjects 1 and 2 initially converge to a norm, then subject 1 is replaced by 3 for enough trials and then 2 is replaced by 4. The final group consisting of totally new subjects 3 and 4 will conform to the norm established by 1 and 2.[13]

Solomon Asch expanded on these results to an even simpler task. Subjects are brought into a lab and sit with others to judge the length of lines such as those in Figure 1. The subject hears the judgment of the others and then makes his own. For several trials, this is a very boring task as it is pretty obvious which line is longer. But then there is a twist. On one of the trials, the first person makes a wrong choice. Then the second person makes the same wrong choice. And so it continues until it is the subject's turn. In Asch's experiment, there were 5 to 12 "conformity" trials out of 10 to 18 total trials. What Asch found was stunning. Between 50 to 80 percent of the subjects yielded to the erroneous majority at least once. Of course, as Asch notes, it is not the subjects perception of the line length (unlike perhaps in the Sherif experiment) that is altered. Many (not all) are simply willing to conform in their behavior.

Other experiments suggest that individuals may conform strongly to their roles (Aronson, Steele, Salinas and

Lustina 1998). A modern day version of this is seen in recent work on stereotype threat. In one early and particularly clever study, African-American and White subjects in the United States are asked to take the GRE. In one condition, they are asked to fill out a questionnaire indicating their sex, major of study and other demographic variables (but not race). In another condition, they are asked to also fill in their race. This simple manipulation—by evoking the race of the person—elicited conformity to the stereotype. The African-American students, who are often stereotyped as less intelligent, responded by fulfilling it. In the condition where race was salient they did far worse than the Whites. In the condition where race was not salient, however, they did exactly the same as the Whites.

Hoff and Pandey (2003) have recently performed a similar experiment on caste in India. Children of lower and upper caste are asked to solve mazes on a piece rate basis. In some cases, caste is made highly salient (through public announcement of the child's caste). When this occurs, the low caste children solve 25 percent fewer mazes. They go on to provide some evidence for a mechanism in this case. When asked to accept or reject a gamble in which there is *no* scope for judgment by an experimenter, making caste salient does *not* produce a caste gap. In stead it is in the case where there is scope for subjective judgment by others that caste appears to have an effect. This suggests that one of the reasons people fall so easily into caste roles is that they expect others to treat them according to these roles.

As Hoff and Pandey note, these type of findings can be helpful for understanding why institutions and inequalities can persist. Norms and institutions can shape what people think is possible. They can shape people's perceptions of how others will respond to them and thereby drive behavior. For example, a lower cast child may feel strongly the norms and stereotypes that go along with being lower caste. This can serve as a powerful deterrent to him to get educated or to seek a higher station in life. In this way, inequalities when defined by well-identified groups can persist.

Policies attempting to reduce inequalities need to be highly cognizant of the prevailing norms. In the low caste case, for example, simply giving supply side incentives or reservations alone may not solve the problem. The tug of the prevailing norms can be stronger then the drives of material interests. The flip side of this logic produces a classic "big push" type argument. If some small group of discriminated individuals does manage to break the norms and succeed, the effect can be powerful. They can serve as role models for many others and remove the norm-induced barrier at least. In these models, the key question is how to promote this initial change and how to then publicize their success.

[13] Camerer and Weber (2003) present an interesting examination of how such norms can arise and evolve over time.

Self Serving Bias and Evaluation

Hastorf and Cantril (1954) got two groups of students, one from Princeton and one from Dartmouth, to watch film of a Princeton-Dartmouth football game. Each student was asked to count the number of penalties committed by both teams. Though both are watching the exact same tape, the counts show that they "saw a different game". Dartmouth students saw an equal number of flagrant and mild penalties committed by both teams. Princeton students, on the other hand, counted three times as many flagrant penalties by Dartmouth as by Princeton and the same number of mild penalties. This experiment illustrates an often-repeated finding in psychology, that beliefs and the perceptions that feed into forming them can be biased. In this case, the students' personal affiliations with the school influenced what they saw. In other cases, it might be prior beliefs or a desire for a particular outcome that might bias perceptions and opinions.

Babcock and Loewenstein (1997) provide a particularly stunning example of this bias. Subjects are asked to bargain over how to deal with a particular tort case (that is based on a real trial that occurred in Texas). Each subject is assigned the role of lawyer for the defendant or plaintiff. They read all the case materials and then bargain with each other over a settlement amount. If they fail to settle, the award will be what the judge decided in the actual case (which is unknown to the subjects at the time of bargaining). Interestingly, subjects are paid as a function of how much they manage to get in the settlement and pay a cost if they go to the judge without settling. Subjects are also asked to assess (in private) how much they think the judge will award them. Finally, some pairs of subjects read the entire description of the case *before* knowing what role they'll play. Others read it afterwards. This order of reading has a large effect. Those who read first settle at 94% rate without going to the judge. But those who read afterwards only settle at 72% rate. Moreover, as a rule, those who read before tended to exaggerate how much the judge would favor them. In short, they exhibit quite biased beliefs based on the position they are in. Plaintiffs think the merits of the case support a large award whereas defendants think it merits a small one. These conflicting beliefs are generated through nothing more than the roles they were assigned. When they read through the case, they selectively interpreted the information they saw in light of their role. Note that this goes against their material interests in one important way. They must pay to go to court yet their biased beliefs get them to go much more often. Much like with the football game, these subjects saw very different cases. In some sense, they saw what they "wanted" to see.

Of all the evidence I've presented, I feel this has the most far-reaching impact for how development policy is practiced and that is why I end with it. I feel this evidence tells us something very important for how development policy ought to be evaluated. A useful example is in the study of Cabot's intervention program for delinquent youth in Cambridge and Somerville, Massachusetts (Powers and Whitmer 1951). This intervention combined all the best available tools at the time for helping these delinquent youth, from tutoring, psychiatric attention to interventions in family conflicts. Those involved in the program raved of its success. They all had very positive impressions. What made the program unique however was that a true

random assignment procedure was used to assign the students. When this data was examined, contrary to the very positive (and likely heartfelt impressions of the caseworkers), there was little effect of the program.

Ross and Nisbett (1991) cite another interesting example: a meta-analysis by Grace, Muench and Chalmers (1966) who study all medical research on the "portacaval shunt". This was a popular treatment for cirrhosis of the liver and 51 studies examined the efficacy of this treatment. The doctors and scientists doing these studies all had the same good intent: to see whether this procedure worked. But the studied differed in one important way: 15 of them used controls but without randomization while 4 of them used truly randomized strategies. Thirteen of the fifteen non-randomized studies were markedly or moderately enthusiastic about the procedure. Yet only one of the randomized studies was markedly or moderately enthusiastic.

What is going on here? I feel the good intentions of the doctors and scientists get in the way. There is always subjectivity in non-randomized trials, what controls to include, what controls not to include, which specification to run, etc. Such subjectivity gives room for self-serving bias to rear its head. Exactly because the researchers on these topics are well intentioned, exactly because they hope the procedure works, it is all to easy for them to find a positive result. Much as with the Dartmouth and Princeton students, they saw in some sense what they wanted to see.

As I noted earlier, I feel both of these examples highlight a very important fact about evaluation. Especially in the development context where most working with a project would like to see it succeed, it is all too easy for self-serving bias to affect evaluations. Beyond the obvious econometric benefits of randomized evaluation, I think this is one of the greatest practical benefits. Randomized trials are a way of minimizing (though obviously not eliminating) researcher's latent biases. They allow us to escape the dangers of biased perception which researchers or field workers are no freer of than anyone else in the population.

Concluding Observations

Much of recent development economics has stressed the importance of institutions. Property rights must be enforced to give appropriate incentives for investment. Government workers must be given appropriate incentives to ensure high quality public service provision. Banking may need to be privatized to ensure a well-functioning credit system that in turn allows for better savings and smoother consumption. The common theme here is that institutions must be improved to help to resolve issues between people. They may reduce externalities, solve asymmetries of information, or help resolve coordination problems. This focus on institutions resolving problems *between* people, rather than *within* a person is natural to economists. The predominant economic model of human behavior leaves little room for individuals themselves to make mistakes. In fact, economists assume that people are unbounded in their cognitive abilities, unbounded in their willpower and unbounded in their self-interest (Mullainathan and Thaler 2001). Once we admit human complexities, institutional design in development becomes not just about solving problems *between* people. It is also about developing institutions that help any one person deal with their own "problems". I hope the small set of examples I have presented illustrate how a deeper understanding of the psychology of people might eventually improve development policy.

References

Akerlof, George A, 1991, "Procrastination and Obedience," *American Economic Review*, 81 (2) pp. 1-19.

Ainslie G., 1992. *Picoeconomics*. Cambridge: Cambridge University Press.

Ardener, Shirley, and Sandra Burman, eds., 1995. *Money-Go-Rounds: The Importance of Rotating Savings and Credit Associations for Women*. Washington DC: Berg Publishers.

Aronson, J., Steele, C. M., Salinas, M. F. Lustina, M. J., 1998, "The effects of stereotype threat on the standardized test performance of college students," in E. Aronson, (Ed.), *Readings About the Social Animal* (8th edition). New York: Freeman.

Ariely, D., and Klaus Wertenbroch, 2002, "Procrastination, Deadlines and Performance: Self-Control by Precommitment," *Psychological Science*, 13, pp. 219-224.
Becker, G., 1993, *Human Capital*. 3rd ed. Chicago, IL: University of Chicago Press.

Asharf, N., Dean Karlan and Wesley Yin, 2004, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product," mimeo, Princeton University.
http://www.wws.princeton.edu/~dkarlan/downloads/SEED.pdf

Babcock, L., and Loewenstein, G., 1997, "Explaining bargaining impasse: the role of self-serving biases." *Journal of Economic Perspectives*, 11, 109-126.

Bertrand, M., Sendhil Mullainathan and Eldar Shafir, 2004, "Behavioral Economics of the Poor," *American Economic Review*, May.

Camerer, C. and Roberto Weber, 2003, "Cultural conflict and merger failure: An experimental approach," *Management Science*, 49(4), 400-415.

Darley, J.M., and B. Latané, 1968, "Bystander intervention in emergencies: Diffusion of responsibility," *Journal of Personality and Social Psychology*, 8, pp. 377-383.

Farkas, Steve and Jean Johnson, 1997. *Miles to Go: A Status Report on Americans' Plans for Retirement.* New York, NY: Public Agenda.

Frederick, S., George Loewenstein, G. and Ted O'Donoghue, (2002), "Time Discounting and Time Preference A Critical Review," *Journal of Economic Literature*, 40(2), 351-401.

Gilbert, D. T., R. W. Tafarodi and P. S. Malone, 1993, "You can't not believe everything you read," *Journal of Personality and Social Psychology*, 65, 221-233.

Gruber, Jonathan and Mullainathan, Sendhil. "Do Cigarette Taxes Make Smokers Happier," mimeo, MIT, March 2002. http://econ-www.mit.edu/faculty/gruberj/files/happy81.pdf

Gugerty, Mary Kay, 2001, "You Can't Saving Alone: Testing Theories of Rotating Savings and Credit Organizations," mimeo, Harvard University.

Hastorf, A. and H. Cantril, "They saw a game: A case study," *Journal of Abnormal and Social Psychology*, 49, pp. 129-134.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R., 2002, "Cooperation, Reciprocity and Punishment: Experiments from 15 small-scale societies." Book manuscript.

Hoff, K. and Priyanka Pandey, 2004, "Belief Systems and Durable Inequalities: An Experimental Test of the Effect of Indian Caste on Performance," World Bank Working Paper.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler, 1990, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 98, pp.1325-1348.

Laibson D., 1997 Golden Eggs and Hyperbolic Discounting. Quarterly Journal of Economics. 62, 443-478.

Madrian, B. C. and D. F. Shea, 2001, "The power of suggestion: Inertia in 401(k) participiation and savings behavior," *Quarterly Journal of Economics*, 116(4), 1149-1187.

Messick, D. M., & Sentis, K. P. (1979). "Fairness and preference," *Journal of Experimental Social Psychology*, l5, 4l8-434.
Mullainathan, S. and Richard Thaler, 2001, "Behavioral Economics," *International Encyclopedia of Social Sciences.* Pergamon Press, 1st edition, October: 1094-1100.

Genesove, David and Chris Mayer, 2001, "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics* 116(4), 1233-1260.

Grace, N. D., H. Muench, and T. C. Chalmers, 1966, "The Present Status of Shunts for Portal Hyperextension in Cirrhosis," *Gastroenterology*, 50, pp. 684-691.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze (1982) "An Experimental Analysis of Ultimatum Bargaining," *Journal of Economic Behavior and Organization*, 3:4 (December), 367-388.

Jacobs, R.C, and D.T. Campbell, 1961, "The perpetuation of an arbitrary tradition through several generations of a laboratory microculture," *Journal of Abnormal and Social Psychology*, 34, pp. 385-393.

Odean, Terry, "Are Investors Reluctant to Realize Their Losses?" 1998, *Journal of Finance*, pp.1775-1798.

O'Donoghue T, and Matthew Rabin, 2001, "Choice and Procrastination," *Quarterly Journal of Economics*. 66, 121-160.

Powers, E., and H. Whitmer. 1951. *An experiment in the prevention of delinquency*. New York: Columbia.

Read, D., George Loewenstein,. And S. Kalyanaraman, 1999, "Mixing virtue and vice: Combining the immediacy effect and the diversification heuristic," *Journal of Behavioral Decision Making*, 12, 257-273.

Regan, R.T., 1971, "Effects of a favor and liking on compliance," *Journal of Experimental Social Psychology*, 7, 627-639.

Ross, L. and Richard E. Nisbett, 1991. *The Person and The Situation: Perspectives on Social Psychology*. McGraw Hill.

Samuelson, W., and Richard Zeckhauser, 1988, "Status quo bias in decision making." *Journal of Risk & Uncertainty*, v. 1, pp. 7-59.

Sherif, M., (1937), "An experimental approach to the study of attitudes," *Sociaometry*, 1, pp.90-98.

Strotz, R. (1956). "Myopia and Inconsistency in Dynamic Utility Maximization." *Review of Economic Studies*, 23, pp. 165-180.

Thaler, R., "Some Empirical Evidence on Dynamic Inconsistency," *Economic Letters*, 8: 201-207.

Thaler R., 1988, "The Ultimatum Game," *Journal of Economic Perspectives.*

Thaler, R., and Shlomo Bernatzi, "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, forthcoming.

Trope, Y., and A. Fishbach, "Counteractive self-control in overcoming temptation," *Journal of Personality and Social Psychology*, 79, pp. 493-506.

Vermeesch, Christel, 2003," School Meals, Educational Achievement and School Competition: Evidence from a Randomized Evaluation," working paper, University of Oxford. http://hicks.nuff.ox.ac.uk/users/vermeersch/schoolmeals.pdf

Wartenbroch, K., 1998, "Consumption self-control by rationing purchase quantities of virtue and vice," *Marketing Science*, 17, pp. 317-337.