

An ESRC Research Group



**Intrinsic motivations on the development frontline:
Do they exist? Do they endure?**

GPRG-WPS-033

**Abigail Barr, Magnus Lindelow,
Jose Garcia-Montalvo and Pieter Serneels**

Global Poverty Research Group

Website: <http://www.gprg.org/>

The support of the Economic and Social Research Council (ESRC) is gratefully acknowledged. The work was part of the programme of the ESRC Global Poverty Research Group.

**Intrinsic motivations on the development frontline:
Do they exist? Do they endure?**

Abigail Barr ^{a 1}
Magnus Lindelow ^b
Jose Garcia-Montalvo ^c
Pieter Serneels ^a

^a *Centre for Study of African Economies, University of Oxford*

^b *Development Economics Research Group, The World Bank*

^c *University of Pompeu Fabra, Barcelona*

December, 2005

Acknowledgements: The experiments and survey upon which this study is based was funded by the World Bank. Abigail Barr's ongoing involvement in the project is funded by the Economic and Social Research Council through the Global Poverty Research Group. We would like to thank Ritva Reinikka and Agnes Soucat for their support and Danila Serra and Daniel Ali for their invaluable work in the field. The findings interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

¹ Corresponding author. Email: abigail.barr@economics.ox.ac.uk Address: CSAE, Department of Economics, University of Oxford, Manor Road, Oxford, OX1 3UQ, United Kingdom.

Intrinsic motivations on the development frontline: Do they exist? Do they endure?

Abstract

Are new recruits to the development frontline intrinsically motivated in a way that may prevent them from becoming unproductive or corrupt? And are they likely to remain thus motivated as their careers progress? We seek answers to these questions using both survey and experimental data relating to a sample of Ethiopian nursing and medical students. We find that, according to four, arguably salient measures, the majority of the students are intrinsically motivated. We also find evidence that intrinsic motivations are socially rather than individually determined, may change as individuals' social contexts change and may be eroded by exposure to an environment in which unproductive behaviour is endemic.

JEL classification: C93 - Field Experiments; C91 - Laboratory, Individual Behavior

Intrinsic motivations on the development frontline: Do they exist? Do they endure?

1. Introduction

Every year thousands of young people join what the World Bank (2003) refers to as the frontline in development and poverty alleviation. This frontline includes the many professionals delivering health, education and other services aimed at helping the poor to help themselves. Achievement of the Millennium Development Goals, a set of targets relating to improvements in income, health, and education for poor people, depends critically on the performance of these professionals. And, in turn, their performance depends on their institutional environment, i.e., factors such as whether and how they are monitored and rewarded for their performance, and on their levels of intrinsic motivation, i.e., on how much satisfaction they derive directly from their work and whether they associate unproductive or corrupt behaviour with direct psychological costs in the form of guilt. All other things being equal, changing their institutional environment so as to strengthen the links between performance and pay or reduce the net expected benefits associated with corrupt or unproductive behaviour will improve service delivery (e.g., Klitgaard, 1988; Bardhan, 1997; Mookherjee and Png, 1995; Chand and Moene, 1999; Becker and Stigler, 1997; Besley and McLaren 1993). However, all other things may not be equal. It may be that when extrinsic incentives are strengthened through improvements in formal institutional design, intrinsic motivations are crowded out (Deci, 1975; Kreps, 1997; Tandler, 1997; Franco, Bennett, and Kanfer, 2002). So, for example, service delivering professionals may perform better when trusted to do so rather than when they are monitored.

Accounting for negative relationships between extrinsic incentives and intrinsic motivations greatly complicates the already difficult task of institutional reform. But, if we ignore these relationships, new reforms may be doomed to failure. To establish whether crowding out should be taken into account we need answers to three questions. First, is crowding out commonplace? And second, are crowding out effects large enough to be of concern? The empirical evidence on crowding out is primarily experimental (e.g. Bohnet, Frey, and Huck, 2001; Cardenas, Stranlund, and Willis, 2000; Fehr, and List, 2004; Fehr and Rockenbach, 2003) and, hence, subject to concerns about external validity. A few quantitative analyses of historical cases of crowding out have been conducted, but none of these relate to the provision of services in developing countries (e.g. Frey and Oberholzer-Gee, 1997; Gneezy Rustichini, 2000a, 2000b). While potentially very difficult, field studies based on either survey data or survey and experimental data combined could yield high returns. However, before embarking on such an exercise, we need to address the third question – is anything there to be crowded out: are professionals on the development frontline intrinsically motivated in the first place? We know of only one paper that touches

upon this issue; Svensson and Reinikka (2004) found that, in Uganda, health care professionals working for religious not-for-profit organizations get paid less but perform better than those in equivalent government positions and interpreted this as evidence of greater levels of intrinsic motivation among the former. While this may, indeed, be the correct interpretation, many institutional factors vary between the two sectors and could be driving all or part of the difference. Further, the study provides us with no insight into levels of intrinsic motivations and how they vary across individuals; we cannot tell, for example, whether those working in the public sector, i.e., the sector upon which institutional reform is most likely to be visited, are intrinsically motivated at all. Here, we address these issues by endeavouring to measure intrinsic motivations directly.

Our investigation focuses on a sample young Ethiopians who, at the time of the study, were months away from joining the development frontline as government-employed nurses and doctors. We know from previous research into health sectors throughout the developing world (Ferrinho, van Lerberghe, Julien, Fresta, Gomes, Dias, Goncalves, and Backstrom, 1998; Ensor and Witter, 2001; McPake, Asiimwe, Mwesigye, Ofumbi, Orthenblad, Streefland, and Turinde, 1999; Killingsworth, Hossain, Hedrick-Wong, Thomas, Rahman, and Begum, 1999; Ferrinho, Omar, Fernandes, Blaise, Bugalho, and van Lerberghe, 2004) and the Ethiopian health sector in particular (Lindelov, Serneels and Lemma; 2005) that the extrinsic incentives that these youngsters will face are weak. Their pay will be low. It is unlikely that their performance, whether good or bad, will ever be assessed or appropriately rewarded. And the way in which they will be allocated jobs by the government will take no account of their preferences or abilities. Further, due to these weaknesses, they may be working alongside colleagues who are frequently absent, possibly because they are maintaining portfolios of jobs, accepting side payments in exchange for preferential treatment (or, in extreme cases, before they are willing to offer any treatment at all) and expropriating public resources for private use or sale. The question we seek to answer is whether these youngsters have any internalized behavioural tendencies or intrinsic motivations that will prevent them from becoming involved in such unproductive and corrupt practices themselves.

Using two distinct methodologies, we construct four measures of intrinsic motivation. Two of these are derived from the youngsters' responses to survey questions, one relating to their reasons for pursuing a career in health and the other to their attitudes towards allowing oneself to be paid for more hours than worked. Both of these measures appear highly salient to the issue at hand. All other things being equal, we would expect those who are pursuing a career in health for altruistic reasons such as a desire to help the needy or poor to work harder and those who express greater disapproval at the idea of accepting overpayment to be less inclined towards opportunistic or corrupt behaviour. However, while apparently salient, these measures are likely to be subject to response bias. Response bias emerges because survey respondents can choose their responses with reference to an objective other than presenting the truth. For

example, if respondents seek the approval of the surveyors either for its own sake or because they think it will lead to other benefits, they may choose responses that further that aim. Unfortunately, both of the questions just described have more and less socially acceptable responses and, this being the case, response bias is highly likely. For this reason, we derive two other measures of intrinsic motivation using economic experiments. In these economic experiments, the youngsters had to forgo significant, real financial payoffs in order to signal that they were anything other than opportunistic. Thus, we derive measures of the youngsters' trust responsiveness and preferences for distributive equity.

All four measures indicate that, at the time of our study, the youngsters were intrinsically motivated, although to varying degrees. The variations across youngsters are not only interesting in their own right but also render it possible to, first, conduct a combined analysis of the measures, and second, address the question of whether the youngsters' intrinsic motivations are likely to endure as their careers progress. Here we acknowledge the conceptual link between intrinsic motivations and the socially transmitted behavioural rules that are commonly referred to as norms (Kreps, 1997). Norms can be seen as benchmarks against which individuals assess their own behaviour. When an individual who has internalized a particular norm engages in behaviour that deviates from that norm they suffer feelings of guilt and are, thus, intrinsically motivated to act in accordance with the norm. This connection between intrinsic motivations and norms implies that the former may be socially rather than individually determined and that a change in social environment may lead to a change in an individual's intrinsic motivations.¹ The youngsters in our sample are about to enter a work environment in which many act in a manner that is inconsistent with high levels of intrinsic motivation. Will they become less intrinsically motivated as a result? We explore the general stability of our four proxies for intrinsic motivations across social contexts by making use of the fact that the youngsters in our sample are distributed across eleven different colleges. We investigate the specific effects of exposure to work in the Ethiopian health sector by making use of the fact that some of the youngsters have already been thus exposed.

The paper has six sections. Following this introduction, section 2 briefly describes the Ethiopian youngsters upon which our study is focused and then presents our two survey-based proxies for intrinsic motivations. Section 3 presents the design of the experiments used to generate the other two proxies, describes the resulting experimental data, and explains how the proxies are calculated. In section 4, we conduct a combined analysis of the four measures. In section 5, we investigate the likely durability of the youngsters' intrinsic motivations. Finally, in section 6, we summarize our results and draw some conclusions.

¹ For examples of psychological and sociological theories relating to the impact of social context on individual behavioural tendencies see Benedict (1934), Grusec and Kuczynski (1997), Cavalli-Sforza and Feldman (1981), and Boyd and Richerson (1985).

2. Ethiopian medical and nursing students and two survey-based proxies for their intrinsic motivations

This study focuses on 299 students from eight clinical nursing schools and the three medical faculties situated across Ethiopia. One third of the sample is medical students just months away from starting their first hospital internship and two thirds are senior clinical nurses about to complete their final year of training and take up their first paid postings. This sample represents approximately half of all the newly operative nursing and medical personnel flowing into the Ethiopian health sector in any single year.

Table 1: Student characteristics

	Mean or proportion	Standard deviation
Personal characteristics		
<i>age (years)</i>	22.645	2.927
<i>female</i>	0.385	
<i>ethnicity</i>		
<i>group 1 (a single language group)</i>	0.268	
<i>group 2 (a single language group)</i>	0.351	
<i>group 3 (4 language groups, same region)</i>	0.184	
<i>group 4 (all remaining groups)</i>	0.197	
<i>parental household expenditure (estimate, in 10,000 Birr p.a.)</i>	5.142	0.598
<i>father has primary school</i>	0.538	
<i>private secondary school</i>	0.097	
Recent and current context		
<i>experience working in the health sector (years)</i>	0.272	1.330
<i>college attended</i>		
<i>college 1</i>	0.100	
<i>college 5</i>	0.100	
<i>college 7</i>	0.100	
<i>college 2</i>	0.100	
<i>college 3</i>	0.100	
<i>college 4</i>	0.084	
<i>college 6</i>	0.097	
<i>college 8</i>	0.064	
<i>college 9</i>	0.087	
<i>college 10</i>	0.067	
<i>college 11</i>	0.100	
Proxies for intrinsic motivations		
<i>desire to help</i>	0.642	
<i>stated integrity</i>	4.068	1.059
<i>value of distributive equity</i>	0.263	0.162
<i>trust responsiveness</i>	1.033	0.682

Notes: Number of observations is 299 except in the case of *desire to help* and *stated integrity* for which there are 296 and 294 respectively.

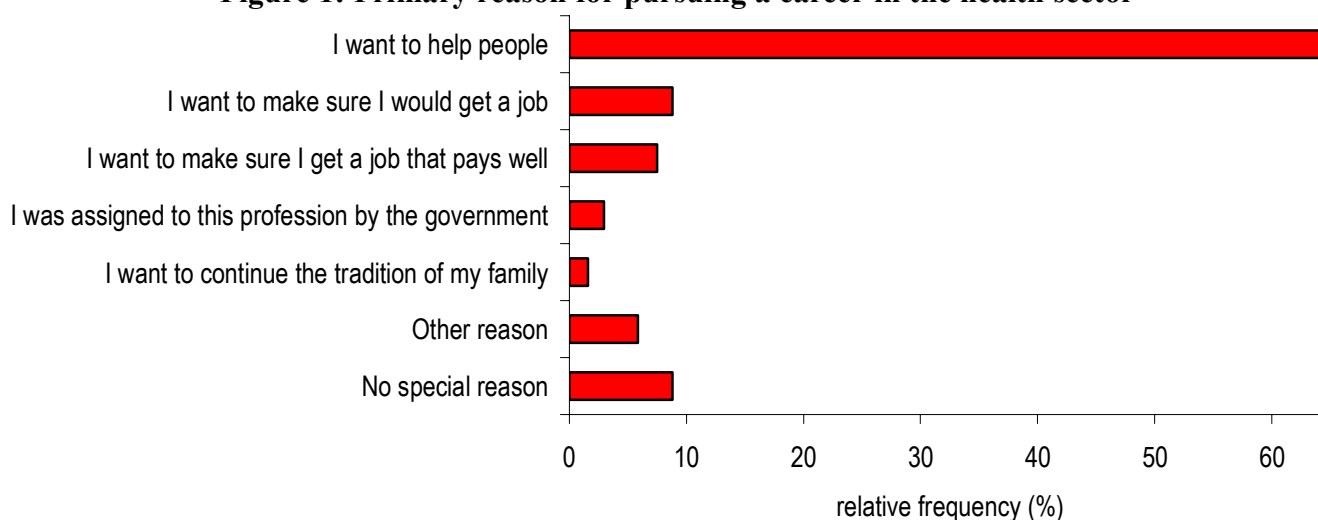
Table 1 presents summary data relating to these 299 students. At the time of the study, their average age was 22.7 years, 38 percent were female, and they came for a wide range of ethnic groups.² Using data

² Note that neither ethnic groups nor colleges are named in the table. This is not an oversight. It follows a deliberate decision on the part of the authors: during our analysis we find evidence of significant behavioural differences across both ethnic groups and colleges. The existence of such differences is salient to our hypothesis testing, while linking behaviours to particular ethnicities or colleges is not and could, in our opinion, be counter productive.

generated by a series of questions about the characteristics of the households in which they grew up, we estimated the annual parental household expenditure for each student. The average for the sample was just over 50,000 Birr (approximately 5,000 USD) per annum, suggesting that they come from relatively well off Ethiopian families. Consistent with this finding, 54 percent of the students' fathers attended at least primary school and nearly ten percent of the students themselves attended private secondary schools.

The questionnaire designed to elicit the information described above also contained a number of questions relating to the students' intrinsic motivations towards their studies and their future careers in the health sector. Here, we focus on the two that we thought would be most salient to their future performance as health workers. First, the students were asked why they had chosen to pursue a career in the health sector. The frequencies with which they gave the five most common answer, some other answer, or reported no special reason for their career choice are presented in Figure 1. Only one of the reasons given, a desire to help people, is indicative of intrinsic motivations. However, this reason was given by over 64 percent of our sample. In the analyses that follow, we use a dummy variable taking the value one for students who said they wanted to help people and zero otherwise as our first measure of intrinsic motivations. Below, we refer to this proxy as the students' *desire to help*.

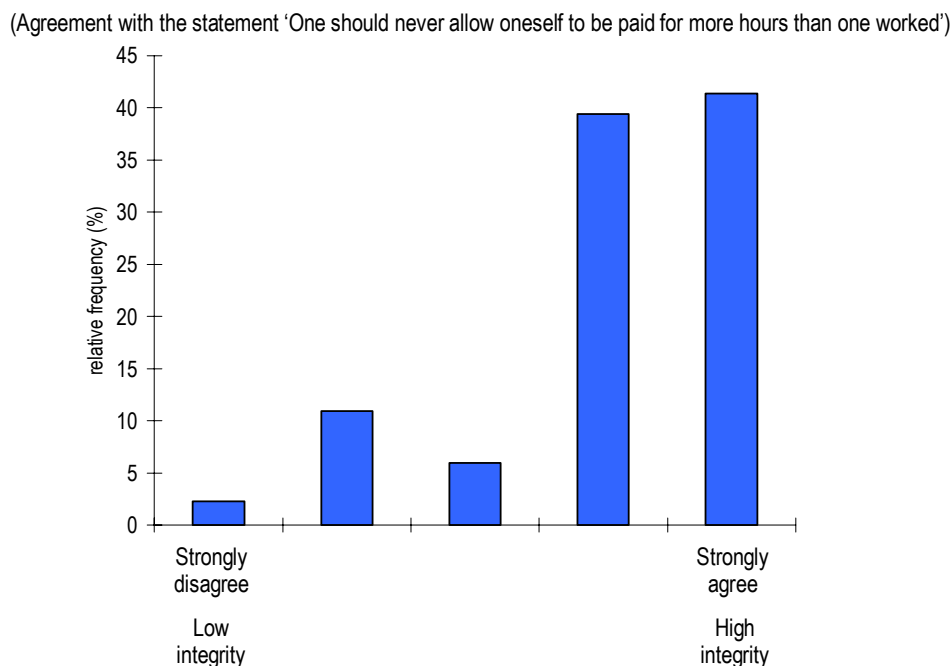
Figure 1: Primary reason for pursuing a career in the health sector



In an endeavour to establish how likely the students are to associate feelings of guilt with underperformance in work, we asked each whether and how strongly they agreed or disagreed with the statement 'One should never allow oneself to be paid for more hours than one worked'. The resulting data, which takes the form of a five point Likart scale (strongly disagree, disagree, neither agree nor disagree, agree, strongly agree) is presented in Figure 2. The figure indicates that 81 percent of the students either agreed or strongly agreed with the statement, while 13 percent disagreed or strongly disagreed. In the analyses that follow, we use this scale in its raw form as our second measure of intrinsic

motivations. Below, we refer to this measure as the students' *stated integrity*. The measure takes values from one, indicating relatively low integrity, to five indicating relatively high integrity.

Figure 2: Variations in stated integrity



Based on these first two, survey-based measures, the students appear highly intrinsically motivated. However, as we mentioned in the introduction, we need to consider the possibility that the students may have answered these survey questions in accordance, not with their intrinsic motivations, but with their beliefs about what may be deemed appropriate. In the case of the second measure, this is equivalent to saying that they told us what the norm is rather than whether and to what extent they associate feelings of guilt with breaking the norm.

3. Experimentally derived proxies for intrinsic motivations

3.1 Experimental design

We asked the youngsters to participate in a Third Party Punishment Game (TPPG) and a Generalized Trust Game (GTG). The TPPG, initially designed by Fehr and Fischbacher (2004), involves three players, Player A, Player B, and Player C. Play is anonymous and one-shot. Player A is given an initial endowment of money and invited to offer some portion (between zero and 100 percent) of that endowment to Player B. Player C is given an initial endowment some of which he can choose to spend on punishing Player A. In Ethiopia, Player A's initial endowment was set at Birr 50 (approximately \$5) and he or she could make offers in multiples of 5. Player C's initial endowment was set at Birr 25 and he or

she could spend Birr 5 to reduce Player A's final payoff by Birr 15 or Birr 10 to reduce Player A's final payoff by Birr 30. Player As made their decisions by dividing their initial endowment between two envelopes. Player C had to fill out a form stating what they would do in the case of each possible division that could be made by Player A, before seeing the actual division made.

The GTG, initially designed by Buchan, Crosson, and Dawes (2002), is similar to Berg, Dickhaut and McCabe's (1995) investment game, in which a Proposer decides how much (zero to 100 percent) of her initial cash endowment to send to a Responder. The amount she sends is tripled by the experimenter before being passed on to the Responder who also receives the same initial endowment. Then, the Responder is invited to send some portion (zero to 100 percent) of the tripled amount back to the Proposer. The GTG differs from this game in one respect: the Responders are sending money back, not to the Proposer who sent to them, but to some other Proposer present in the same session. Buchan et al. (2002) found that both trust and reciprocity were lower in the GTG than in the standard BDM game. In Ethiopia, the Proposers and Responders' initial endowments were set at Birr 40 and each player could send multiples of Birr 10. Proposers made their decision by dividing their initial endowment between two envelopes. Responders had to fill out a form stating what they would do in the case of each possible amount sent by their Proposer, before seeing the actual amount sent.

We chose the TPPG because we wished to measure the value that each of the youngsters in our sample place on distributive justice or equity. This being the case, we were particularly interested in how they would behave in the role of Player C (the finer). So, we recruited additional youngsters from the same colleges to assume the roles of Players A and B. (For these we have no survey data.) We chose the GTG because we wished to measure our survey respondents' tendencies to pass on resources that have been entrusted to them. Compared to Responders in the more commonly played BDM investment game, the decision of Responders in the GTG is conceptually closer to the decisions made on a day-to-day basis by public servants; they receive resources, which they are expected to pass on to others. So here, we wanted the youngsters in our sample to assume the role of Responder. So, we retained half of the additional youngsters recruited for the TPPG to assume the Proposer role in the GTG.

Each student attended the one or one of the two experimental sessions held at their own college. Where two sessions were held, steps were taken to ensure that participants in the first session could not inform participants in the second about the games. A total of 20 sessions were held, with between 29 and 60 students attending each.³ At the start of every session, our survey respondents were shown into one classroom and the other students into a second classroom. In each classroom, the students sat at amply

³ All 20 sessions were led by the same team comprised of one graduate student from Oxford University and two Ethiopian research assistants. One research assistant was assigned to each room, and the graduate student remained in the first classroom throughout. The research assistants delivered the verbal instructions in Amharic.

spaced desks. No talking was allowed. The TPPG was explained verbally and each student was given a table showing what their earnings would be under all possible scenarios. The Player Bs in the second classroom received envelopes containing slips of paper informing them of their role and explaining that they would receive their earnings and the end of the session. The Player As each received two envelopes one containing their endowment and one into which they could place the amount of money they wished to pass on to Player B. The Player Cs, i.e., our survey respondents in the first classroom, after being taught the game and receiving the tables of earnings, were asked to detail their strategies on the forms provided. The envelopes and forms were processed at a desk in the front of the first classroom.

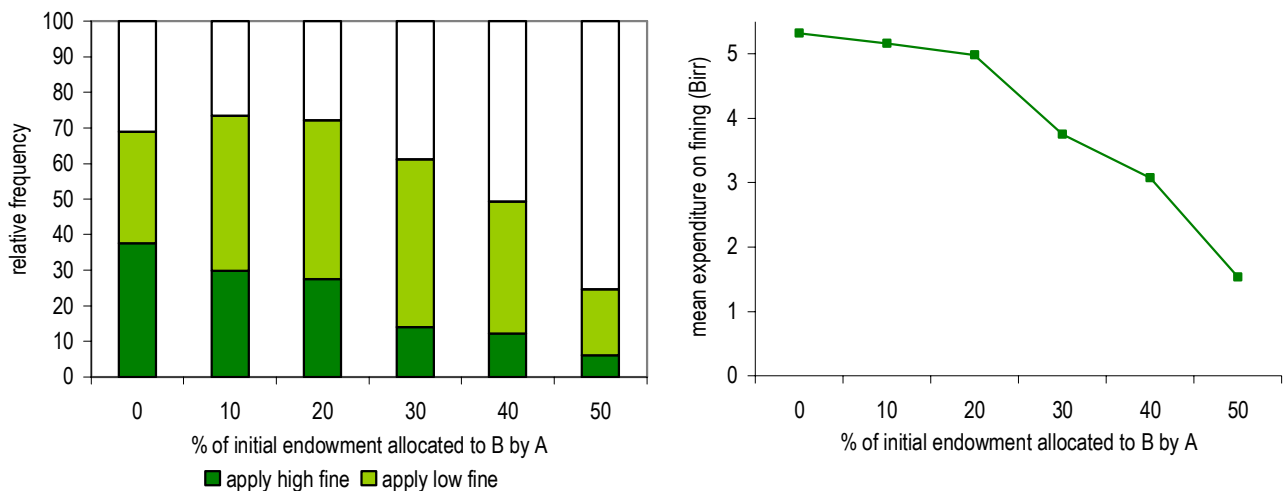
Before knowing the outcome of the TPPG, the students were taught the GTG. Again, the instructions were given verbally and each student was given a table of earnings under all possible scenarios. In the second classroom the students who had been passive Player Bs in the TPPG received two envelopes, one containing a slip of paper informing them that they were Proposers and their initial endowment of Birr 40, the other to be used to send whatever amount they wished to a Responder. The remaining students in the second classroom received envelopes containing slips of paper informing them that they had not been picked to participate in the GTG but would receive compensation for being asked to sit quietly until the end of the experimental session. In the first classroom, the Responders were, again, asked to detail their strategies on the form provided. The envelopes and forms were, once again, processed at a desk in the front of the first classroom.

Before looking at the data generated by these experiments, it is worth reflecting a little longer on the nature of these experiments and why they were chosen. Note, first, that neither of the experiments was designed to directly identify crowding out effects as this was not our objective. Nor were they designed to simulate environments in which individuals tend to be unproductive or corrupt in the manner we see in Abbink and Renner (2002), Abbink (2004), Azfar and Nelson (2002), Barr, Lindelow and Serneels (2004), and Cameron, Chaudhuri, Erkal, Ganadharan (2005). This design choice was harder. Had we involved the youngsters in a highly contextualized experiment, we might have generated apparently more salient experimental measures. However, at the same time we might have restricted their salience to those real world situations that most closely resemble the experiment. And we might also have framed our entire endeavour in the eyes of the students as one relating to seeing how corrupt they may become as time progresses. As we plan to revisit the youngsters as their careers progress with the objective of investigating not only their performance on the job, but also the problems and constraints that they see themselves facing and the strategies they devise to cope in variably difficult environments, we preferred simpler and more abstract experimental designs.

3.2 Experimental data and the construction of two proxies for intrinsic motivations

The fining strategies chosen by the students are summarized in Figure 3. In the right and left hand panels, respectively, the frequencies with which the youngsters imposed the high and low fines and their average expenditure on fining are plotted against six of the possible allocations that Player A could make to Player B. Around 70 percent of the youngsters chose to fine Player As who allocated 20 percent or less of their initial endowment to Player B. As allocations to Player B increased from 20 to 50 percent, the proportion of youngsters choosing to fine steadily declined and the average amount spent on fining declined monotonically as allocations to Player B increased from zero to 50 percent.⁴ These data indicate that the majority of the students value distributive equity to such a degree that they are willing to forgo some portion of their expected payoff in order to express their disapproval when Player As act selfishly. However, the data also indicate that there is considerable variation in their evaluations.

Figure 3: Fining strategies in the TPPG



We can use each student's fining strategy to construct a measure of the value they place on distributive equity equal to the amount by which they would reduce their expenditure on fining following a one Birr increase in the amount that Player A allocates to Player B. To construct this proxy we fit a linear regression model with a fixed intercept indicating no fining in the case of the equitable division to each youngster's fining strategy over the range of allocations depicted in Figure 3. Because the intercept is fixed, each regression yields a single estimated slope coefficient. This slope coefficient is our proxy.⁵ Figure 4 presents a histogram of the estimated slope coefficients and clearly displays the extent to which the students vary in terms of their *valuation of distributive equity*. Around 14 percent appear not to value

⁴ A significant proportion of the youngsters fined Player As making the equitable division. However, in the majority of cases, these fines are symptomatic of problems understanding the game (Barr *et al.* 2005). Barr *et al.* (2005) also presents a more detailed and thorough analysis of the fining data.

⁵ Any particular summarizing method will work better when applied to some strategies than others. In other words, the summarizing method introduces a variable (across respondents) amount of measurement error into the proxies. This notwithstanding, of the estimated coefficients: 58.90% are significant at the 1% level; 73.79% are significant at the 5% level; and 79.94% are significant at the 10% level.

it at all. Just over one quarter of the sample falls in the modal range. A one Birr increase in the amount that Player A allocates to Player B leads to a 0.2 to 0.3 Birr reduction in these students' expenditures on fining. Around 41 percent of the sample value distributive equity to this degree. Just under 20 percent of the sample place a positive value on distributive equity that is below this modal range, while over 40 percent place a value on distributive equity that is above this modal range. On average, a one Birr increase in the amount that a Player A allocates to Player B caused the students to reduce their expenditure on fining by 0.27 Birr.

Figure 4: Individual valuations of distributive equity

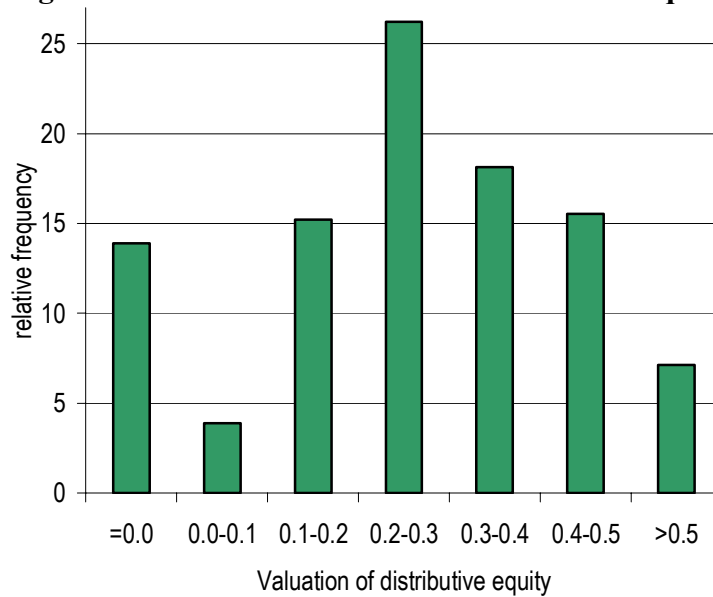
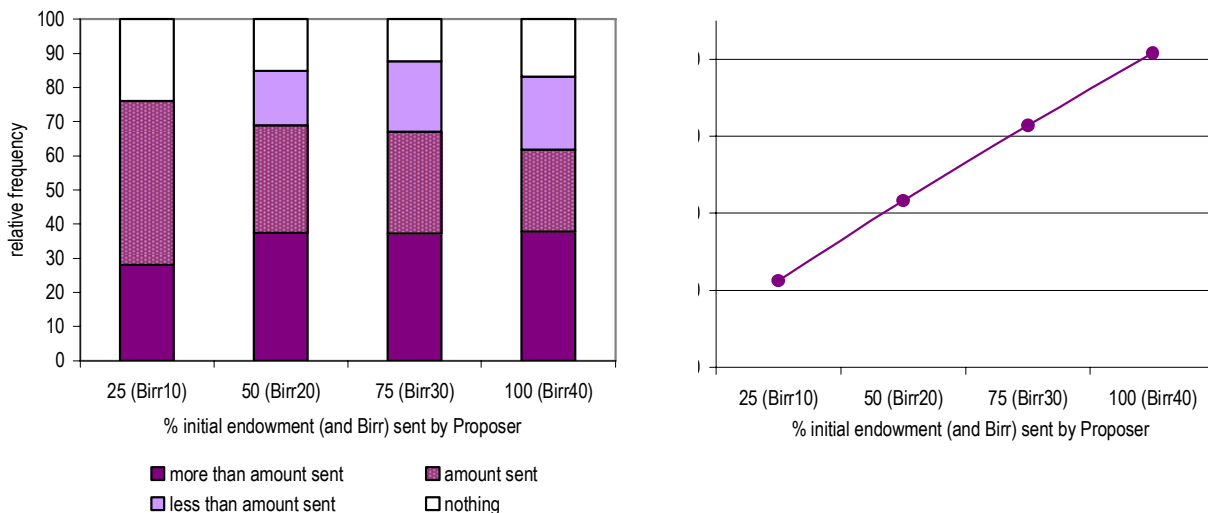


Figure 5: Strategies in the GTG

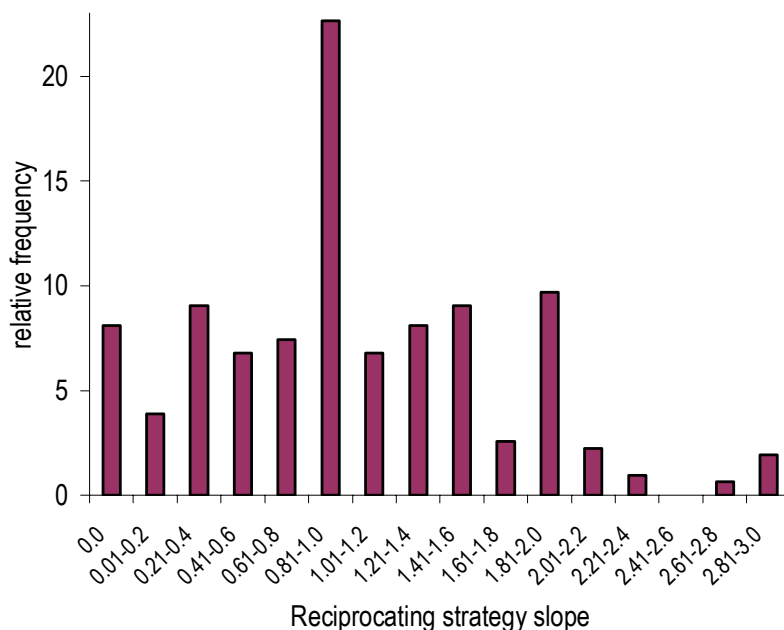


The strategies chosen by the students in the GTG are summarized in Figure 5. In the left hand panel of this figure, the frequencies with which the youngsters chose to return more than the amount sent to them by a Proposer, exactly the amount sent to them by a Proposer, a positive amount, less than that sent to

them by a Proposer, and nothing are plotted against the four possible amounts that the Proposers could send. The right hand panel plots the mean amounts returned by the students given each of the four possible amounts sent by the Proposers. At least 75 percent of the students chose to return a positive amount in the case of each possible amount sent by the Proposers. On average, the students chose to return slightly more than the amount sent, while keeping almost two thirds of the tripled amount for themselves. However, as with the fining strategies, there is evidence of considerable variation.

We can use each student's strategy in the GTG to construct a measure of their *trust responsiveness*, i.e., how much more they would send back to the Proposers given a one Birr increase in the amount that a Proposer sends to them. To do this we fit a linear regression model with a fixed intercept indicating that nothing is returned when nothing is sent.⁶ Once again, because the intercept is fixed, each regression yields a single estimated slope coefficient. This estimated slope coefficient is our second proxy relating to the youngsters' social preferences. Figure 6 presents a histogram of these estimated slope coefficients and clearly displays the extent to which the students vary with respect to their *trust responsiveness*. The distribution has a single mode, relating to a slope of one, i.e. of returning the amount entrusted, while keeping all additional returns on the trusting act. 22 percent of the sample falls at this mode, while the remainder are fairly evenly distributed across the full range of possible levels of trust responsiveness (zero to three). On average, a one Birr increase in the amount Proposers send is associated with a 1.03 Birr increase in the amount the students would return.

Figure 6: Individual trust responsiveness



⁶ In this case, the analysis of Barr *et al.* (2005) indicates that the summarizing method we have chosen is highly appropriate. However, this notwithstanding, the summarizing may have introduced a variable (across respondents) amount of measurement error into the proxies. Of the estimated slope coefficients: 58.90% are significant at the 1% level; 80.26% are significant at the 5% level; and 86.73% are significant at the 10% level.

4. Comparing the measures

Recall that our four measures of intrinsic motivation have different strengths and weaknesses. In particular, the two survey measures have high apparent salience but may be subject to response bias, while the two experimental measures are likely to be free of response bias but are not so apparently salient. Especially with this being the case, we can learn a great deal by looking at whether and to what extent the measures are correlated with one another.

Table 3 presents the correlation coefficients between all the possible proxy pairings.⁷ *Stated integrity*, the *value of distributive justice*, and *trust responsiveness* are all positively and significantly correlated with each other, whereas *the desire to help* proxy is not significantly correlated with any of the others and shows a slight tendency towards being negatively associated.

Table 2: correlations between intrinsic motivation proxies

	<i>Desire to help</i>	<i>Stated integrity</i>	<i>Value of distributive equity</i>
<i>Stated integrity</i>			
Correlation coefficient	-0.080		
P-value (H ₀ : no correlation)	0.165		
<i>Valuation of distributive equity</i>			
Correlation coefficient	-0.047	0.126 **	
P-value (H ₀ : no correlation)	0.417	0.028	
<i>Trust responsiveness</i>			
Correlation coefficient	-0.009	0.113 **	0.227 ***
P-value (H ₀ : no correlation)	0.875	0.049	0.000

Notes: * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

The correlations between the experimental variables and the *stated integrity* could be viewed as evidence that the former is not entirely driven by response bias and the latter have some salience to the issue of corruptibility. The apparent uniqueness of the *desire to help* variable is, however, troubling. It raises questions (although, in our opinion, does not answer them) about the sincerity of the youngsters' statements about wishing to help people.

⁷ Where a correlation is not the ideal test of correspondence, chi-squared and other test have been conducted. In all cases, these other tests concur with the correlations and yield similar levels of significance. We report the correlations because they allow the reader to see the sign and significance of the relationship at a glance.

5. The durability of intrinsic motivations

Now we turn to the durability of intrinsic motivations. We conduct four statistical regressions each taking one of the proxies for intrinsic motivations as its dependent variable. The set of explanatory variables in each model is the same and can be divided into two subsets, one relating to the students' identity and family background, and one relating to their current and recent context. In each regression, significant coefficients on the subset of explanatory variables capturing differences in the students' current and recent context may be taken as evidence of instability in intrinsic motivations as long as we are prepared to accept that those coefficients are not biased due to omitted variables and selection processes.

The subset of variables relating to the students' identity and family background is included in order to minimise omitted variable and selectivity bias in the coefficients of primary interest. It contains eight variables: a dummy variable indicating that the respondent's *father has primary school*; an estimate of their annual *parental household expenditure*; a dummy variable indicating whether they attended a *private secondary school* as opposed to a government funded secondary school;⁸ and three *ethnic dummy variables*; their *age* in years; and a dummy variable indicating whether they are *female*.⁹ Of these variables, the first two capture the affluence of the childhood home, the third captures educational quality, the ethnic dummies capture differences in the wider cultural contexts in which the respondents received their primary socialization, and the age and gender variables capture any differences over cohorts and across genders in the primary socialization process and the effects of aging on intrinsic motivations that do not relate to individual-specific changes in context.

The subset of variables relating to the students' recent and current contexts includes their years of *experience working in the health sector*, and ten *college dummy variables*. The former captures effects that exposure to the behavioural norms prevailing in the Ethiopian health sector has on individual intrinsic motivations, while the latter captures the effects of college-specific norms. In the case of the latter, it is important to bear in mind that the dummies will also capture any differences in levels of intrinsic motivations between medical and nursing students. The means, proportions, and, where appropriate, standard deviations for all the variables across the sample of 299 youngsters for whom we have complete or near complete data are presented in Table 1.

⁸ With few exceptions all of the students attained the same number of years of schooling, leaving aged 18. This is why we do not include years of schooling in the regressions.

⁹ Earlier runs of the analysis contained a number of additional variables, including a set of religion dummies, a set of dummies capturing father's occupation, a variable capturing mother's education, the students' heights, and a dummy indicating whether their father is still alive. None of these were significant and many displayed a high degree of multicollinearity with other variables (e.g. fathers' occupation was closely related to their level of education).

Table 3: Regression analyses of social preferences

	<i>Desire to help</i>		<i>Stated integrity</i>		<i>Value of distributive equity</i>		<i>Trust responsiveness</i>	
	(probit estimation)		(ordered probit estimation)		(weighted least squares estimation)		(weighted least squares estimation)	
	dF/dx ⁺	s.e	Coeff.	s.e	Coeff.	s.e	Coeff.	s.e
Childhood context								
<i>father has primary school</i>	-0.039	0.064	-0.107	0.152	-0.014	0.028	0.068	0.116
<i>parental household expenditure</i>	-0.015	0.056	-0.083	0.122	0.016	0.024	0.124	0.103
<i>private secondary school</i>	-0.011	0.013	0.237	0.299	0.076	0.034 **	0.118	0.136
<i>age</i>	0.091	0.068	0.053	0.031 *	0.009	0.006	0.026	0.023
<i>female</i>	-0.042	0.105	0.144	0.160	0.014	0.031	0.134	0.140
<i>ethnicity (basis of comp. is group 4)</i>								
<i>group 1 (dummy variable)</i>	-0.080	0.091	-0.280	0.180	-0.010	0.037	-0.308	0.179 *
<i>group 2 (dummy variable)</i>	-0.125	0.083	0.095	0.194	-0.023	0.036	-0.288	0.142 **
<i>group 3 (dummy variable)</i>	-0.015	0.097	-0.199	0.211	-0.016	0.042	-0.511	0.175 ***
Recent and current context								
<i>experience working in the health sector</i>	0.043	0.029	0.009	0.061	0.000	0.010	-0.056	0.025 **
<i>college attended (basis of comp. is college 1)</i>								
<i>college 2 (medical, dummy variable)</i>	-0.160	0.135	0.661	0.265 **	-0.006	0.051	0.498	0.181 ***
<i>college 3 (medical, dummy variable)</i>	0.035	0.118	0.536	0.301 *	0.028	0.044	0.310	0.159 *
<i>college 4 (dummy variable)</i>	0.141	0.113	-0.231	0.292	-0.060	0.060	-0.019	0.318
<i>college 5 (dummy variable)</i>	-0.084	0.135	0.199	0.274	-0.030	0.055	-0.122	0.180
<i>college 6 (dummy variable)</i>	0.195	0.098	-0.238	0.296	-0.098	0.062	-0.216	0.254
<i>college 7 (dummy variable)</i>	0.010	0.130	0.715	0.277 **	-0.065	0.052	-0.070	0.235
<i>college 8 (dummy variable)</i>	0.257	0.093 *	0.082	0.364	-0.016	0.077	0.372	0.314
<i>college 9 (dummy variable)</i>	0.174	0.105	0.869	0.304 ***	-0.004	0.057	0.545	0.265 **
<i>college 10 (dummy variable)</i>	0.222	0.102	0.854	0.341 **	-0.109	0.062 *	0.411	0.264
<i>college 11 (dummy variable)</i>	-0.097	0.131	0.624	0.271 **	-0.029	0.052	0.029	0.159
<i>constant</i>	-	-	-	-	-0.009	0.180	-0.141	0.721
Number of observations	296		294		299		299	
R-squared	0.112		0.065		0.093		0.196	
Joint sig. of ethnic dummy variables	0.426		0.155		0.931		0.033	
Joint sig. of college dummy variables	0.020		0.001		0.517		0.004	
Joint sig. of diffs. between medical colleges	0.278		0.276		0.723		0.019	
Joint sig. of diffs. between nursing colleges	0.027		0.001		0.709		0.093	

Notes: ⁺ In Probit regression, for dummy variables, dF/dx is for discrete change. * significant at 10% level; ** significant at 5% level; *** significant at 1% level. All standard errors are heteroscedasticity-robust. In regressions for *value of distributive justice* and *trust responsiveness* observations are weighted by one over the estimated standard error of the dependent variable to take account of errors due to subject cognition and the methods used to summarize subjects' behavior in the experiments.

Table 3 presents the regressions. The first column contains a probit analysis of the students' *desire to help*. The second column contains an ordered probit analysis of their *stated integrity*. And the third and fourth columns contain linear regression analyses of the students' *valuations of distributive equity* and *trust responsiveness* respectively.¹⁰ In the latter two, the observations are weighted to take account individual-specific errors in the measurement of the dependent variable.¹¹ In all the regressions, the

¹⁰ Additional regression analyses (not presented) indicated that the censoring of these two dependent variables, due to the design of the experiments, does not impact significantly on the estimated coefficients.

¹¹ The weights are set equal to one over the square of the standard error associated with each student's *valuation of distributive equity* and *trust responsiveness* estimates. These standard errors are generated along with the proxies during the regression analyses described in section 3 above.

standard errors have been adjusted to take account of interdependences within sessions and heteroscedasticity associated with the other explanatory variables.

The regression in the first column indicates that the students' *desire to help* is not associated with their family background or identity. Neither is it associated with *experience working in the health sector*. However, among nurses, it is associated with the college that they attend. The regression in the second column indicates that, among the first subset of explanatory variables, only their *age* significantly affects their *stated integrity*. The youngsters' *experience working in the health sector* has no significant effect. However, *stated integrity* varies significantly across the full set of eleven colleges, the set of three medical colleges, and the set of eight nursing colleges. The regression in the third column indicates that only one aspect of the students' family background and identity, their attendance at a *private secondary school*, affects their *valuation of distributive equity*. *Experience working in the health sector* has no significant effect and, while one of the college dummies is significant, neither the full set of dummies nor the subsets relating to medical and nursing colleges separately are jointly significant. Attendance at a private secondary school is associated with an increase in the value placed on distributive equity that is equivalent in size to one third of the sample mean valuation. The regression in the fourth column indicates that, with respect to their identity and family background, only the students' *ethnicity* significantly affects their *trust responsiveness*. *Trust responsiveness* varies significantly across the full set of eleven colleges, the set of three medical colleges, and the set of eight nursing colleges and is significantly and negatively associated with *experience working in the health sector*. An additional year of experience leads to a decline in trust responsiveness equivalent in size to one half of the mean level of responsiveness.

Before concluding this section, we need to consider the possibility that the results of primary interest reported above are due to some form of selection process, or reversed causality. The significance of the college dummies in three of these regressions could be due to a selection process rather than a causal link from college attended to individual intrinsic motivations. However, this would seem unlikely. The majority of the students in our sample are attending a college in either the region in which they are born or the region associated with their ethnicity. Now, this correspondence between ethnicity, geography and college might well play a role in determining the levels of intrinsic motivations prevailing in particular colleges. However, we controlled for ethnicity in the analysis we present and controlled for birthplace in earlier runs of the analysis finding that it did not change the results of interest and led to problems of multicollinearity with the ethnicity variables. The other factor that is likely to affect selection is the students' abilities and the reputations of the colleges for producing well trained doctors and nurses. We have no data relating to the apparent academic abilities of the students prior to their entry into the colleges. However, we did ask the students to sit tests of their knowledge of medical theory and practice

and found that among the nursing students one fifth of the variation in their scores was explained by the schools they attended, suggesting that academic standards do vary across the colleges. Further, these test scores are neither significant nor reduce the significance of the college dummies when entered into the regressions taking intrinsic motivations as the dependent variable. Finally, it seems unlikely that the causal link between experience working in the health sector and *trust responsiveness* runs from latter to former rather than the other way around.

6. Summary and conclusions

Our aim in this paper was to establish whether new Ethiopian nurses and doctors are intrinsically motivated in a way that may prevent them from becoming unproductive or corrupt and whether they are likely to remain intrinsically motivated as their careers in the health sector progress. Our interest in these questions stemmed from a concern that, if such intrinsic motivations exist but are not taken into account during institutional reforms, the latter could lead to a worsening rather than an improvement in the provision of health services to the poor.

We constructed four measures each relating to a different type or dimension of intrinsic motivation. According to these measures, nearly two thirds of our sample of medical and nursing students chose their careers out of a desire to help people, a similar proportion placed a positive value on distributive equity, over 80 percent expressed disapproval with the idea of accepting only pay commensurate with hours worked, and over 90 percent showed some degree of trust responsiveness. However, levels of intrinsic motivation varied markedly between students.

With respect to whether the students' intrinsic motivations are likely to endure once they start their careers as salaried health workers in the Ethiopian public sector, the evidence was mixed. We found that the students' *desire to help people*, their *integrity*, and their *trust responsiveness* varied significantly depending on the college they attended and argued, with the help of various statistical tests, that this should be taken as evidence of the instability of individual intrinsic motivations across social context. Further, *trust responsiveness* appeared to be rapidly eroded by exposure to the Ethiopian health sector.

These results suggest that new recruits to the development frontline are intrinsically motivated in ways that may prevent them from becoming unproductive or corrupt. However, their levels of intrinsic motivation may be vulnerable to erosion through exposure to the prevailing work and related social environment.

Taken at face value, these results suggest that any improvements in service delivery to the poor will be dependent on the strengthening of the formal institutional environment and, hence, the extrinsic incentives faced by service delivering professionals and that such improvements are unlikely to lead to reductions in performance due to the crowding out of intrinsic motivations. However, it is important to bear in mind the many weaknesses in our analysis that we mention above. We view this research as a first exploratory step towards a quantitative analysis of the role of intrinsic motivations in service delivery to the poor. In future analyses, researchers might usefully consider alternative dimensions and measures for intrinsic motivation. In making their choices of proxies, researchers should bear in mind that only one of our survey-based measures was corroborated by our experimentally derived measures. In our future research we plan to make use of data on the actual performance and corrupt behaviour of the youngsters once they have entered paid work, as well as information on the extrinsic incentives they face and certain key events during their first months in the workplace that might have significantly affected their levels of motivation.

References

- Abbink, K., B. Irlenbusch and E. Renner (2002) 'An Experimental Bribery Game' *Journal of Law, Economics, and Organization*, 18: 428-54.
- Abbink, K. (2004). 'Staff rotation: A powerful weapon against corruption?', *European Journal of Political Economy*, forthcoming.
- Akerlof, G. A. (1980) 'A theory of social custom, of which unemployment may be one consequence' *The Quarterly Journal of Economics*, 94: 749-75.
- Akerlof, G. A. (1982) 'Labor Contracts as Partial Gift Exchange' *Quarterly Journal of Economics*, 97: 543-69.
- Azfar O. and R. Nelson, (2002), 'Transparency, Wages and the Separation of Powers: An Experimental Analysis of the Causes of Corruption', IRIS working paper.
- Bardhan, P. (1997) 'Corruption and development: A review of issues' *Journal of Economic Literature*, XXXV, September: 1320-1346.
- Barr, A., Lindelow M. and P. Serneels (2004) 'To Serve the Community or Oneself: The Public Servant's Dilemma', CSAE WPS/2003-11.
- Barr, A., Garcia-Montalvo J. Lindelow M. and P. Serneels (2005) 'Strategy Choice and Cognitive Ability in Field Experiments', GPRG WPS/2005-?
- Becker, G. S. and G. J. Stigler (1997) 'Law Enforcement, Malfeasance, and Compensation of Enforcers' *Journal of Legal Studies*, 1-19.
- Benedict, R. (1934) *Patterns of culture*. Boston, MA: Houghton Mifflin.
- Besley, T. and J. McLaren (1993) 'Taxes and bribery: the role of wage incentives' *Economic Journal*, 103(416): 119-141.

- Bewley, T. F. (1999). Why Wages Don't Fall During a Recession?
- Bohnet, I., B. Frey, and S. Huck (2001) 'More Order with Less Law: On Contractual Enforcement, Trust, and Crowding' *American Political Science Review*, 95:1, pp. 131-44.
- Cardenas, J. C., J. Stranlund, and C. Willis (2000) 'Local Environmental Control and Institutional Crowding-out' *World Development*, 28:10, pp. 1719-33.
- Boyd, R. and P. J. Richerson (1985) *Culture and Evolutionary process*. Chicago, IL: University of Chicago Press.
- Cameron, L., A. Chaudhuri, N. Erkal, L. Ganadharan (2005) 'Do attitudes towards corruption differ across cultures? Experimental evidence from Australia, India, Indonesia, and Singapore' Working paper, University of Melbourne.
- Cavalli-Sforza, L. L. and M.W. Feldman (1981) *Cultural transmission and evolution*. Princeton, NJ: Princeton University Press.
- Chand, S. K. and K. O. Moene (1999) 'Controlling fiscal corruption' *World Development* 27(7): 1129-40.
- Ensor, T. and S. Witter (2001) 'Health economics in low income countries: adapting to the reality of the unofficial economy' *Health Policy*, 57(1): 1-13.
- Fehr, E. and S. Gaechter (2002) 'Do Incentive Contracts Crowd Out Voluntary Cooperation?' Mimeo, University of Zurich.
- Fehr, E., G. Kirchsteiger, and A. Riedl (1998) 'Gift Exchange and Reciprocity in Competitive Experimental Markets' *European Economic Review*, 42: 1-34.
- Fehr, E. and J. List (2004) 'The hidden costs and returns of incentives: Trust and trustworthiness among CEOs' *Journal of The European Economic Association*.
- Fehr, E. and B. Rockenbach (2003) 'Detrimental effects of sanctions on human altruism' *Nature*, 422:13 March, pp. 137-40.
- Frey, B. and F. Oberholzer-Gee (1997) 'The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out' *American Economic Review*, 87:4, pp. 746-55.
- Ferrinho, P., W. Van Lerberghe, M. R. Julien, E. Fresta, A. Gomes, F. Dias, A. Goncalves, and B. Backstrom (1998) 'How and why public sector doctors engage in private practice in Portuguese-speaking African countries' *Health Policy Plan*, 13(3): 332-8.
- Ferrinho, P., M. Omar, M. Fernandes, P. Blaise, A. Bugalho, W. van Lerberghe (2004) 'Pilfering for survival: How health workers use access to drugs as a coping strategy' *Human Resources for Health* 2(1): 4.
- Franco, L. M., S. Bennett, and R. Kanfer (2002) 'Health sector reform and public sector health worker motivation: a conceptual framework' *Social Science and Medicine*, 54: 1255-66.
- Gneezy, U. and A. Rustichini (2000a) 'A Fine is a Price' *Journal of Legal Studies*, 29:1, pp. 1-17.
- Gneezy, U. and A. Rustichini (2000b) 'Pay enough or don't pay at all' *Quarterly Journal of Economics*, 115:2, pp. 791-810.

- Gray, C. W. and D. Kaufmann (1998) 'Corruption and Development' *Finance and Development*, March, 1998: 7-10.
- Grusec, J.E. and L. Kuczynski (1997). *Parenting and children's internalisation of values: A handbook of contemporary theory*. New York: John Wiley & Sons.
- Gupta, S., H. Davoodi and R. Alonso-Terme (1998) 'Does Corruption Affect Income Inequality and Poverty?' IMF Working Paper 98/76.
- Klitgaard, R. (1988) *Controlling Corruption*. Berkeley: University of California Press.
- Kreps, D. M. (1997) 'Intrinsic Motivation and Extrinsic Incentives', *American Economic Review*, 87(2), 359-64.
- Killingsworth, J., N. Hossain, Y. Hedrick-Wong, S. Thomas, A. Rahman, T. Begum (1999) 'Unofficial fees in Bangladesh: price, equity and institutional issues' *Health Policy Plan*, 14(2): 152-63.
- McPake, B., A. Asimwe, F. Mwesigye, M. Ofumbi, L. Orthenblad, P. Streefland, and A. Turinde (1999) 'Informal economic activities of public health workers in Uganda: implications for quality and accessibility of care' *Social Science and Medicine*, 49(4): 849-65.
- Mookherjee, D. and I.P.L. Png (1995) 'Corruptible law enforcers: how should they be compensated' *Economic Journal*, 105, January: 149-159.
- Nye, J. S. (1967) 'Corruption and Political Development' *American Political Science Review*, 61(2): 417-27.
- Svensson, J. and R. Reinikka (2004) 'Working for God? Evaluating service delivery of (religious) not-for-profit health facilities in Uganda', *CEPR Working Paper* No. 4214.
- Tanzi V. and H. Davoodi (1997) 'Corruption, Public Investment, and Growth' IMF Working Paper WP/97/139.
- Tendler, J., (1997) *Good Governance in the Tropics*, London: The Johns Hopkins Press.