Essays on Econometric Evaluation of Education Outcomes in Developing Countries

Inaugural-Dissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München
Volkswirtschaftliche Fakultät

2008

vorgelegt von
Johannes Moritz Metzler

Referent: Prof. Dr. Ludger Wößmann
Korreferent: Prof. Dr. Joachim Winter
Promotionsabschlussberatung: 04. Februar 2009

Acknowledgements

I am indebted to my supervisor Prof. Ludger Wößmann who sparked my interest in the economics of education and enabled timely completion of this dissertation through valuable guidance; in a short period of time I was able to gain a lot from his knowledge in this discipline, especially during collaboration for the third chapter. I am also very grateful to Prof. Joachim Winter who advanced my work through extensive feedback on all chapters. Both managed to distill the essence of their areas of expertise for me. I thank Prof. Martin Kocher, too, who kindly agreed to be my third examiner.

I am grateful to the participants of Prof. Winter's and Prof. Wößmann's internal seminars whose productive comments and suggestions helped me advance more than once; special thanks go to Oliver Falck, Florian Heiss, and Guido Schwerdt. I am also thankful for the support received during my enriching experience at the World Bank, especially by my supervisor and co-author of the fifth chapter, Gabriel Demombynes. Thanks also to Lorena Alcazar, Omar Arias, Monserat Bustelo, Gustavo Canavire, Raji Jayaraman, Jann Lay, Maria-Beatriz Orlando, Anne Pillay, and Jaime Saavedra.

The stimulating environment of the Munich Graduate School of Economics made this dissertation possible in the first place. Foremost, I thank the German Research Foundation for a full scholarship. Furthermore, I thank Mrs. Buchmayr and Mrs. Szantone-Sturm for administrative support, and Prof. Rady for his supervision of the program.

I am deeply indebted to the friends and colleagues who have accompanied me throughout this journey, especially to Moritz Hahn and Martin Watzinger who endured innumerable discussions about my work and pushed me constantly by challenging my ideas. I also thank Benno Bühler, Karolina Kaiser, Veronika Scharfenberg and Sebastian Doedens for enriching this time of my life.

Foremost, however, my greatest gratitude goes to Anna for being an inspiration and drive, and to my parents and my brother for their love and support without whom none of my achievements would have been possible.

Johannes Metzler
Munich, September 2008

Table of Contents

List of Tables

List of Figures

Chapter 1

1 The Economics of Education in Developing Countries

"Literacy is a bridge from misery to hope. It is a tool for daily life in modern society. It is a bulwark against poverty, and a building block of development, an essential complement to investments in roads, dams, clinics and factories. [...] For everyone, everywhere, literacy is, along with education in general, a basic human right. Literacy is, finally, the road to human progress and the means through which every man, woman and child can realize his or her full potential."

Kofi Annan

"We must go on fighting for basic education for all, but also emphasize the importance of the content of education."

Amartya Sen

This dissertation presents four empirical essays that offer findings at the interface of two disciplines within economics: the economics of education on the one hand and development economics on the other. The economics of education employs economic analysis to examine both the determinants of education and education's impact on individuals and on the economies which they inhabit. Development economics employs economic analysis to examine the development process in low- and middle-income countries. Taken together, the disciplines reveal important insights for development policymakers.

We have known for a long time that education is a crucial driver in the formation of individual and societal well-being. At the micro level, Mincer (1974) popularized the concept of rates of return to schooling, i.e., the earnings premium that is associated with an additional year or level of schooling. These returns have been estimated with micro data in countless studies ever since.[1] The relation between education and earnings also translates to the macro level: higher intra-national education equality results in higher earnings equality (Katz, Autor 2005) and higher levels of education lead to higher growth levels cross-nationally (Hanushek, Wößmann 2008).

Many development economists and policymakers regard education as the most powerful instrument of the poor to escape poverty. It now ranks among the foremost goals on the international development agenda: the UNESCO Education for All (EFA) Initiative and the United Nations Millennium Development Goals (MDGs) both aim to improve education in developing countries comprehensively by 2015, primarily by achieving universal primary education and gender equality at all education levels. The emphasis on education is not solely motivated by education's direct positive effects on individuals' earnings. Broader and more equitable education appears to have positive external effects on society and contributes to the improvement in non-monetary poverty dimensions as well.[2] For example, research has associated higher education levels with increased labor productivity and mobility, better resource management and faster diffusion of information and technologies (Porter 1998). Furthermore, we may expect particularly strong effects of education on the level of health within a country's population (Grossman, Kaestner 1997). Education, particularly of women, in developing countries is associated with family planning and reduced fertility, lower infant and maternal mortality, immunizations and better nutrition (World Bank 1993, 2001). Also, more primary education and gender equality in the educational system appears to be correlated with higher levels of democracy (Barro 1999). Ultimately, education enables people to escape

[1] Studies estimating these so-called "Mincer returns to schooling" have consistently shown the explanatory power of education in wage determination at the individual level, even though this is often not equivalent to the actual causal return from additional schooling (Heckman, Lochner, Todd 2006).

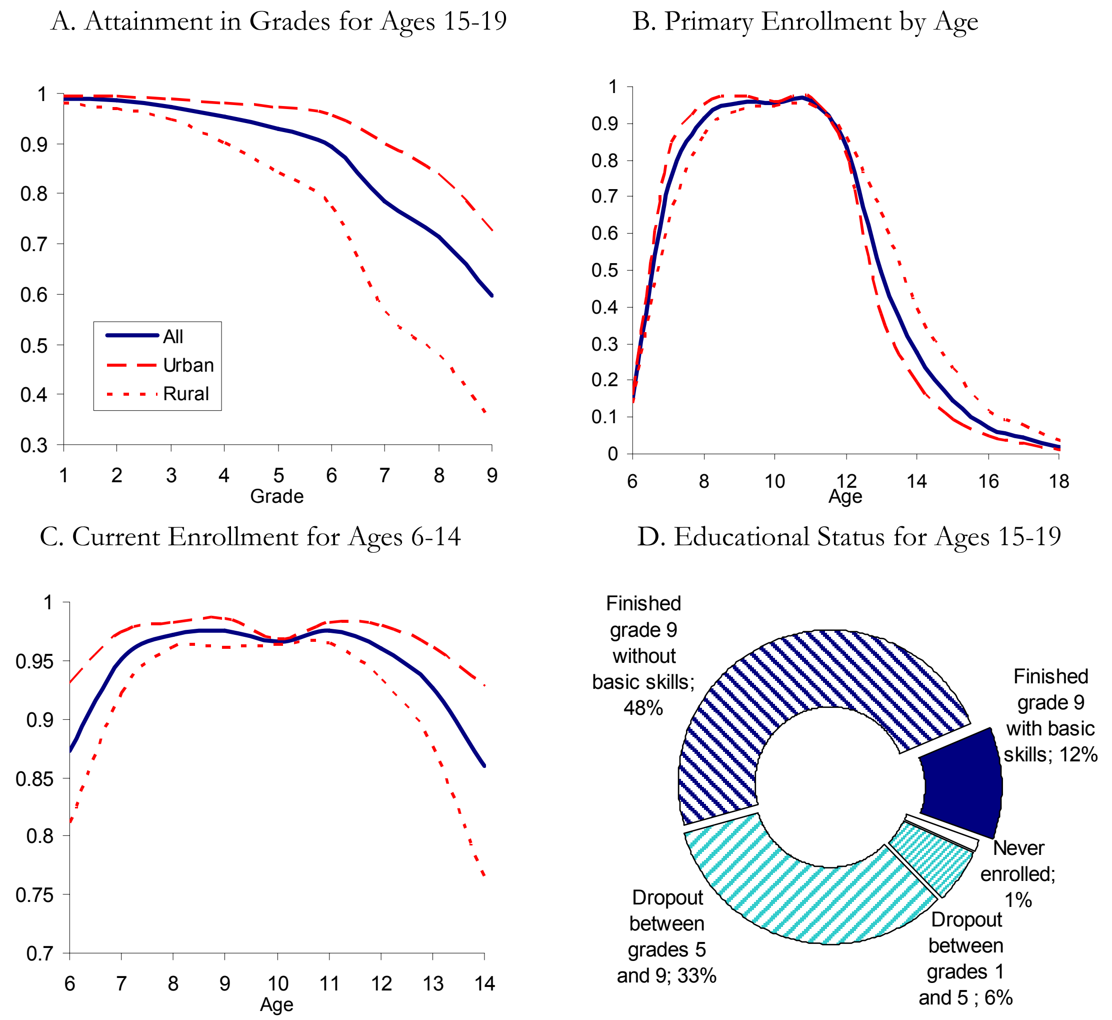
[2] For more comprehensive surveys see, for example, Grossman (2006) or McMahon (2004).

poverty by building up "human capabilities" (Sen 1999), i.e., the capabilities to function as members of society.

Yet, the state of education in many developing countries is still far below an acceptable level. Researchers measure the state of education with quantitative indicators, such as school enrollment and primary completion rates, or qualitative indicators, such as standardized test scores, which provide an indication of cognitive capabilities. Primary enrollment and completion rates have risen substantially in the past two decades in developing countries. Between 1991 and 2005, the net enrolment ratio[3] in the developing world rose from 79 to 86 percent (UNESCO 2007). At the same time, the global average primary completion rate[4] rose from 72 to 77 percent (World Bank 2003). This process has not been uniform, however, with deteriorating figures for some sub-populations, regions and countries. Also, developing countries suffer from educational inefficiency, i.e., children repeating grades and dropping out of school before completing the educational cycle. UNESCO (1998) estimates that the public cost of educational wastage in the world's less developed regions was around 16 percent of public current expenditure in 1995. The rate of survival to the last primary school grade[5] was less than 80 percent in developing countries in 2004 (UNESCO 2007). Many countries will thus fail to achieve the educational MDGs at the current pace (World Bank 2008). In addition, these quantitative measures of education still underestimate the actual gap between the knowledge and skills of children in industrialized and developing countries. Hanushek and Wößmann (2008) estimate that in 14 developing countries with sufficient data on the quantity and quality of schooling, fewer than one third of 15- to 19-year-olds have reached basic literacy in cognitive skills, even though a substantial share of these students has finished the 9th grade.

[3] The net enrolment ratio expresses enrolment of the official age group for a certain level of education as a percentage of the population in that age group.
[4] The primary completion rate measures the share of all children of official graduation age who complete primary school in a given year.
[5] The survival rate is the percentage of a cohort of students who are enrolled in the first grade of an education cycle in a given school year and are expected to reach a specified grade, regardless of repetition.

Figure 1-1. State of Education in Peru, 2000



Source: Panel A-C: calculations by Filmer (2008), based on the Peru Demographic and Health Survey 2000. Panel D adapted from Hanushek, Wößmann (2008).

A poignant example of this pattern can be found in Peru. The first three chapters in this dissertation deal with determinants of Peruvian schooling outcomes. The country is a good case for examining questions of educational policy because it combines the basic characteristics of many developing countries with the above mentioned patterns of quantity and quality of schooling. Peru is typical for a developing country in that it has heterogeneous geographic regions and ethnic groups, a divide between developed urban and sparsely populated rural areas, and persistent poverty (52 percent in 2004) despite positive and stable macroeconomic performance in recent years (World Bank 2005). Peru spends about 3 percent of GDP on education (World Bank 2007). Figure 1-1 displays schooling figures for Peru in 2000. Panel A shows that 90 percent of 15- to 19-year-olds have finished the 6 grades of primary education, but that there is a sharp drop afterwards, especially in rural areas. Consequently, only 60 percent of students finish 9 grades of schooling. There are two primary reasons for this: first, many students, both urban and rural, remain in primary school after age 11 although the primary cycle should be completed between age 6 and 11 (see panel B). For example, at age 14, almost 30 percent of children are still in primary school – this does not count those who have dropped out prematurely. Second, enrollment figures drop sharply after age 11, especially in rural zones (see panel C). While enrollment is close to 95 percent from ages 7 to 11 for both rural and urban areas, it reaches 86 percent at age 14. Taken together, the data imply that the Peruvian school system is very inefficient due to high grade repetition, overage and early drop-out. Not surprisingly, results from the Programme for International Student Assessment (PISA) in 2001 confirm that educational inefficiency in Peru correlates with overall weak educational quality. Hanushek and Wößmann (2008) find that in 2000, only one in five Peruvian students aged 15 to 19 that had finished 9th grade is functionally literate (see panel D). Since only 60 percent of this student cohort had completed 9 grades, only 12 percent of the entire cohort is adequately literate.

Given the low level of education in developing countries, positive externalities can justify policy interventions to reduce private underinvestment in education.[6] Although economists hotly debate the appropriateness of financing education publicly, the rationale for public interventions is strongest in developing countries because externalities are likely to be very large at low levels of literacy (Hanushek 2002). Policy interventions can be made on two related fronts: in-school and out-of-school (Randall, Anderson 1999). In-school factors relate directly to the educational system and the local schooling environment; out-of-school factors originate in the political, economic or social environment.

The case for addressing *in-school* factors by increasing the *quantity* of school inputs is much stronger for developing than developed countries. Educational researchers have traditionally addressed in-school determinants of educational outcomes by focusing on

[6] Other common justifications for public intervention in education are „economies of scale, market failures in general, and redistributive motives" (Hanushek 2002, p. 2064).

quantitative measures of educational inputs (Hanushek 2003). Behind this focus is the belief that educational spending and quality are closely related. As the importance of teachers for students' learning is undisputed, much of the political and academic debate has centered on the influence of the quantity of teachers, i.e., the student-teacher ratios or class sizes. Estimates of class size effects for both the developed and developing world display a great amount of inconsistency, and do not generally confirm that smaller class sizes have cost-effective economic impacts. If at all, there is more support for this effect in developing countries. This phenomenon may indicate that there are heterogeneous effects at different levels of overall spending or diminishing returns to reducing class size (Hanushek 2003).

Overall, econometric problems plague many of the estimates on the student-teacher ratio given that experimental evidence is rare. The use of different non-experimental evaluation strategies like exploiting rule-induced discontinuities (cf. Angrist, Lavy 1999), while necessary, prevents internationally comparable results for developing countries. Economists thus have to derive estimates on a case-by-case basis, depending on data availability and the institutional settings. Even though input-based policies seem to have failed largely, "It does not mean that money and resources *never* matter" (Hanushek 2003, p. F89). Particularly in developing countries, where resources are sometimes very low, some quantitative inputs appear to matter, such as physical facilities (cf. Hanushek 1995), blackboards (Glewwe, Jacoby 1994) or flipcharts (Glewwe et al. 2004).

In addition, we have reason to believe that the *quality* of inputs is a crucial *in-school* determinant of educational success. We know from U.S. evidence that schools and teachers vary greatly in quality, as measured by the variability in learning that they induce among children across classrooms. For example, Rivkin, Hanushek, and Kain (2005) estimate that the difference between an average teacher and one at the 85th percentile of the teacher quality distribution results in a difference of more than 4 percentile rankings in the student test score distribution in a given year. Hanushek (1992) concludes that the difference between a teacher at the 5th and one at the 95th percentile of the teacher quality distribution in a given year is the equivalent of an entire year of learning.

The existing body of research has been unable to explain this heterogeneity using easily observable quantitative measures of inputs, such as class size. Input quality may be a more important determinant of the currently observed achievement differences. Nevertheless, there is little evidence that teacher characteristics that are commonly used as criteria in hiring and compensation policies, such as education, experience or credentials, matter a great deal (Hanushek, Rivkin 2006). There may thus be scope for identifying the effect of more important components of teacher quality, such as teacher academic skills (Eide et al. 2004), which could lead to better human resource policies. Rice (2003) suggests that teacher academic skills may be particularly important for at-risk students. This is an important finding for developing countries where both teachers and students may have very low academic skills.

In addition to in-school factors, *out-of-school* factors play an important role for educational outcomes, particularly in developing countries. One of these factors is health: not only is there a causal effect of education on health, but healthy and well-nourished children are also more efficient producers of human capital (Grossman, Kaestner 1997). Well-identified academic studies have shown the positive effects of early childhood nutritional status as measured by height-for-age on timely enrollment in school, reduced grade repetition, learning outcomes, and school attainment. Similarly, there is strong evidence for positive effects of health interventions such as iron supplementation, deworming treatments and food supplementation on various educational outcomes.[7] Another important out-of-school determinant of education is the economic situation of the household. Income is strongly associated with schooling in the developing world (Filmer, Pritchett 1999). Poorer households tend to have lower preferences for education and lower means to acquire goods that facilitate educational progress, such as textbooks. Also, there is a large body of literature suggesting that income volatility and shocks contribute to child labor, which in turn has strong detrimental effects on grade repetition, drop-out behavior and learning outcomes (cf. Edmonds 2007).

All of this evidence illustrates the importance of well-identified findings from education economics to policymakers in developing countries. Developing countries are poor by definition. To determine whether these countries are using their scarce resources

[7] Compare Glewwe, Miguel (2007) for a review of the studies.

as effectively and efficiently as possible, it is imperative to examine whether current policies produce the intended results. Such an understanding would then help policymakers to identify opportunities for increasing the effectiveness of the educational system. Analyzing the drivers of education outcomes and to develop policy recommendations from these findings is particularly important as part of this process.

One method for developing such policy recommendations is a careful impact evaluation of factors which influence educational outcomes. Impact evaluation is a method of analysis and policy tool that detects the causal effect of an event, program or input on indicators of interest, the so-called treatment effect on the treated (cf. Ravallion 2007). In theory, this effect is identified by comparing the outcome in presence of the intervention with the outcome in absence of the intervention; in practice, the same unit of observation cannot be observed in these two states of nature at the same time. Thus there needs to be some comparison between the treated group and an adequate untreated control group. In most cases, a simple comparison between these two groups without further assumptions yields biased results of this effect. Only in experimental studies, where placement of the program is randomized, does a simple comparison reveal the desired effect magnitude. As a consequence, it is desirable to set up interventions with randomizations and proper control groups ex-ante. As experiments in educational systems are still rare, researchers often resort to non-experimental methods in order to produce analyses ex-post.[8]

Consequently, chapters 2 to 4 of this dissertation employ retrospective data to estimate causal impacts of events or educational inputs on schooling outcomes in Peru. The analyses employ various techniques to overcome the usual problems of endogeneity, i.e., simultaneity and omitted variables. Each of these three chapters addresses one determinant of education outcomes within the areas mentioned above, respectively: the quantity of teachers, the quality of teachers, and the impact of out-of-school factors.

CHAPTER 2 contributes to the debate on input quantities by examining the effect of lower student-teacher ratios in rural Peru where more than one in five enrolled primary students fails class every year. This problem of educational inefficiency in Peru is pressing but poorly understood; we thus evaluate the impact of adding a second teacher to primary

[8] Compare Glewwe and Kremer (2006) for some examples of educational experiments in developing countries.

single-teacher schools on enrollment and grade completion to assess lower class sizes as a potential remedy. Matched difference-in-difference analysis shows a positive enrollment effect of about 13 percent, mainly from reduced between-year drop-out *before* treatment, possibly in anticipation of improved schooling. Grade completion *levels* are increased after treatment due to the enrollment effect; the actual decrease in the student-teacher ratio of about 40 percent, however, does not lead to a further significant improvement in grade completion *rates*. Increasing teacher quantity is thus unlikely to solve Peru's problem of educational inefficiency.

CHAPTER 3 contributes to the debate on input quality by examining whether teacher academic skills determine teacher quality and improve student test scores. In contrast to the second and fourth chapter, we examine student learning instead of school progression as education outcomes since these findings could help to close a more general research gap, beyond the specific problems of Peru. We exploit a unique dataset that provides test scores in two subjects for both students and their teachers to estimate the causal effect of teacher subject knowledge on student achievement using within-teacher within-student variation. By including student- and teacher-fixed effects, our model circumvents biases from omitted variables and selection. The results indicate that a one standard-deviation increase in teacher test scores increases student test scores by about 4 percent of a standard deviation, and by even more when correcting for measurement error. This finding is especially important for the developing world, where both students and teachers often have very low academic skills.

CHAPTER 4 addresses out-of-school determinants of educational success: it contributes to the relatively new literature on the impact of climate change on schooling outcomes, a commonly overlooked aspect in assessing its economic impacts on poverty-related outcomes. We estimate the effect of natural hazard damages to farmland on primary school grade non-completion rates in rural Peru. Since children were enrolled before the disasters occurred, this serves to learn about the costs of getting rural children through school despite changes in their economic environment. The results indicate that a damage of 42 hectares of average farmland, or 18 hectares of subsistence farmland, causes one schoolchild not to complete the grade he or she is enrolled in. The analysis thus predicts that natural hazards account for several hundred yearly cases of grade failure in rural Peru, and that out-of-school factors can be important determinants of educational inefficiency.

Finally, adding to the research on the economic consequences of education referred to at the beginning, CHAPTER 5 provides a perspective on the relation between changing skill premiums at the micro level and earnings inequality at the macro level in Argentina during the 1980s and 1990s. Largely due to the lack of evidence for competing explanations, skill-biased technical change is the most likely explanation for the increases in the returns to education that occurred in Argentina in the 1990s. Using a semi-parametric re-weighting variance decomposition technique, the analysis shows that during the same period there was an increase in the returns to unobserved skill. This finding lends support to the hypothesis that skill-biased technical change may have been a driver of increases in inequality in Argentina. Additionally, the pattern of changes suggests that the growth in returns to unobserved skills may have been partly responsible for the relative deterioration of informal salaried wages during the 1990s.

Even though the findings of chapters 2 to 4 cannot easily be generalized to all developing countries they do fall in line with the existing research surveyed above. In summary, there is not much support for the effectiveness of pure input-quantity based expenditure policies. While some quantitative inputs have been shown to matter selectively (cf. Glewwe, Kremer 2006), the effect in most cases is dubious, even in situations when expenditure levels are very low. In particular, the evidence for strong positive effects of decreased student-teacher ratios is meager (Hanushek 2003). Nevertheless, teacher salaries eat up most of the educational budget in developing countries (World Bank 2007). They are typically determined as a function of supposed teacher quality characteristics, such as education and experience. While we have no strong evidence for these to matter, we still know that teachers differ strongly in the knowledge they convey to children (Hanushek, Rivkin 2006). Teacher subject knowledge is one of the factors contributing to teacher quality in Peru. Furthermore, out-of-school factors, such as weather shocks, play a significant role in the cost-benefit calculation of households; economic hardship can turn them away from school and towards more basic needs of survival (Edmonds 2007). Educational policy alone cannot address this problem because increases in the returns to education probably would have to be enormous to counterbalance the deterioration of economic situations.

In conclusion, there are formidable problems to improving education in developing countries beyond providing access to all children. In a broad sense, most good development policies are good education policies since better economic conditions also allow larger shares of the population to benefit from education. But in a narrow sense, there is much scope to improve educational policy itself in developing countries. For example, introducing proper incentives can be one promising part of the solution in order to align the interests of the suppliers and consumers of education. For example, the World Bank (2001, p. 51) notes with respect to the teaching profession in Peru that there is "a lack of incentives in the system to encourage commitment, professional development, and higher performance, which could translate into better student achievement." In contrast, we know that conditional cash transfer programs can improve school attainment and the economics of poor households by tying benefit payments to the attendance of children in school (cf. Rawlings, Rubio 2003). Teacher performance pay can have strong effects on learning outcomes (cf. Muralidharan, Sundararaman 2008) and possibly attract more productive teachers into the workforce (cf. Falk, Dohmen 2008). This is one promising area for more research, and decision makers need more solid and relevant evidence from impact evaluations of policies in developing countries to close the "evaluation gap" (Center for Global Development 2006) and channel funds towards their most productive use.

References

Angrist, Joshua D., Lavy, Victor (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics, 114*(2), pp. 533–575.

Barro, Robert J. (1999). Determinants of Democracy. *Journal of Political Economy, 107*(6), pp. S158-83.

Center for Global Development (2006). *When Will We Ever Learn? Improving Lives through Impact Evaluation: Report of the Evaluation Gap Working Group.* Washington, DC: Center for Global Development.

Edmonds, Eric V. (2007). Chapter 57 Child Labor. In T. Paul Schultz and John A. Strauss (Ed.), *Handbook of Development Economics.* Volume 4, pp. 3607–3709. Elsevier.

Eide, Eric, Goldhaber, Dan, Brewer, Dominic (2004). The Teacher Labour Market and Teacher Quality. *Oxford Review of Economic Policy, 20*(2), pp. 230–244.

Falk, Armin, Dohmen, Thomas (2008). You Get What You Pay For: Incentives and Selection in the Education System. Program on Education Policy and Governance Working Papers Series No. 08-06.

Filmer, Deon (2008). *Educational Attainment and Enrollment around the World.* Retrieved July 22, 2008, from World Bank: http://econ.worldbank.org/projects/edattain.

Filmer, Deon, Pritchett, Lant (1999). The Effect of Household Wealth on Educational Attainment: Evidence from 35 Countries. *Population and Development Review, 25*(1), pp. 85–120.

Glewwe, Paul, Jacoby, Hanan (1994). Student Achievement and Schooling Choice in Low-income Countries: Evidence from Ghana. *Journal of Human Resources, 29*(3), pp. 843–864.

Glewwe, Paul, Kremer, Michael (2006). Chapter 16 Schools, Teachers, and Education Outcomes in Developing Countries. In Eric A. Hanushek & Finis Welch (Eds.), *Handbook of the Economics of Education.* Volume 2, pp. 945–1017. Elsevier.

Glewwe, Paul, Kremer, Michael, Moulin, Sylvie, Zitzewitz, Eric (2004). Retrospective vs. Prospective Analyses of School Inputs: the Case of Flip Charts in Kenya. *Journal of Development Economics, 74*(1), pp. 251–268.

Glewwe, Paul, Miguel, Edward A. (2007). Chapter 56 The Impact of Child Health and Nutrition on Education in Less Developed Countries. In T. Paul Schultz and John A. Strauss (Ed.), *Handbook of Development Economics.* Volume 4, pp. 3561–3606. Elsevier.

Grossman, Michael (2006). Chapter 10 Education and Nonmarket Outcomes. In Eric A. Hanushek & Finis Welch (Eds.), *Handbook of the Economics of Education.* Volume 1, pp. 577–633. Elsevier.

Grossman, Michael, Kaestner, Robert (1997). Effects of Education on Health. In Jere R. Behrman & Nevzer Stacey (Eds.), *The Social Benefits of Education,* pp. 69–123. Ann Arbor: University of Michigan Press.

Hanushek, Eric A. (1992). The Trade-Off between Child Quantity and Quality. *Journal of Political Economy, 100*(1), pp. 84–117.

Hanushek, Eric A. (1995). Interpreting Recent Research on Schooling in Developing Countries. *World Bank Research Observer, 10*(2), pp. 227–246.

Hanushek, Eric A. (2002). Chapter 30 Publicly Provided Education. In Alan J. Auerbach & Martin Feldstein (Eds.), *Handbook of Public Economics.* Volume 4, pp. 2045–2141. Elsevier.

Hanushek, Eric A. (2003). The Failure of Input-Based Schooling Policies. *Economic Journal, 113*(485), pp. F64-98.

Hanushek, Eric A., Rivkin, Steven G. (2006). Chapter 18 Teacher Quality. In Eric A. Hanushek & Finis Welch (Eds.), *Handbook of the Economics of Education.* Volume 2, pp. 1051–1078. Elsevier.

Hanushek, Eric A., Wößmann, Ludger (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature, 46*(3).

Heckman, James J., Lochner, Lance J., Todd, Petra E. (2006). Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. In Eric A. Hanushek & Finis Welch (Eds.), *Handbook of the Economics of Education.* Volume 1, pp. 307–458. Elsevier.

Katz, Lawrence F., Autor, David H. (1999). Chapter 26 Changes in the Wage Structure and Earnings Inequality. In Orley C. Ashenfelter & David Card (Eds.), *Handbook of Labor Economics.* Volume 3, Part 1, pp. 1463–1555. Elsevier.

McMahon, Walter W. (2004). The Social and External Benefits of Education. In Geraint Johnes & Jill Johnes (Eds.), *International Handbook of the Economics of Education,* pp. 211–259. Surrey: Edward Elgar Publishing.

Mincer, Jacob (1974). *Schooling, Experience, and Earnings* (Vol. 2). New York: National Bureau of Economic Research.

Muralidharan, Karthik, Sundararaman, Venkatesh (2008). Teacher Performance Pay: Experimental Evidence from India. Program on Education Policy and Governance Working Papers Series No. 08-16.

Porter Michael (1998). Microeconomic Foundations of Competitiveness: The Role of Education. In *The Global Competitiveness Report 1998.* Geneva: World Economic Forum.

Randall, Laura, Anderson, Joan B. (1999). *Schooling for Success: Preventing Repetition and Dropout in Latin American Primary Schools.* Armonk, NY: Sharpe.

Ravallion, Martin (2007). Chapter 59 Evaluating Anti-Poverty Programs. In T. Paul Schultz and John A. Strauss (Ed.), *Handbook of Development Economics.* Volume 4, pp. 3787–3846. Elsevier.

Rawlings, Laura B., Rubio, Gloria (2003). Evaluating the Impact of Conditional Cash Transfer Programs: Lessons from Latin America. World Bank Policy Research Working Paper No. 3119.

Rice, Jennifer K. (2003). Teacher Quality: Understanding the Effectiveness of Teacher Attributes. Washington, DC: Economic Policy Institute.

Rivkin, Steven G., Hanushek, Eric A., Kain, John F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica, 73*(2), pp. 417–458.

Sen, Amartya K. (1999). *Development as Freedom.* New York: Knopf.

UNESCO (1998). *Wasted Opportunities: When Schools Fail: Repetition and Drop-out in Primary Schools.* Paris: EFA Forum Secretariat; UNESCO.

UNESCO (2007). *Education for All: Global Monitoring Report 2008*. Oxford: Oxford Univ. Press.

World Bank (1993). *World Development Report: Investing in Health.* Washington, DC.: World Bank.

World Bank (2001). *Engendering Development: Through Gender Equality in Rights, Resources, and Voice.* Oxford: Oxford Univ. Press.

World Bank (2003). *A Chance for Every Child: Achieving Universal Primary Education by 2015.* Washington, DC: World Bank.

World Bank (2005). *Peru. Opportunities for All. Peru Poverty Assessment.* Washington, DC: World Bank.

World Bank (2007). *Toward High-quality Education in Peru: Standards, Accountability, and Capacity Building.* Washington, DC: World Bank.

World Bank (2008). *MDGs and the Environment: Agenda for Inclusive and Sustainable Development.* Washington, DC: World Bank.

Chapter 2

2 Does the Second Teacher Matter? Effects on Enrollment and Grade Completion in Primary Single-Teacher Schools in Rural Peru

2.1 Introduction

In the process towards universal primary education – Millennium Development Goal No. 2 – many developing countries are scaling up their primary school coverage while the quality of the system deteriorates as educational expenditures do not increase alike. In international student achievement tests, some developing countries with high coverage perform dismally. As quality lags behind, the system becomes clogged by students that do not progress through school in time, a phenomenon often termed *educational wastage* (UNESCO 1998). Many students keep repeating the same grades because they are not promoted to higher grades, creating a vicious cycle: over-aged students become a drain on the remaining class by diverting scarce education materials and teacher attention away from others. Also, repeaters are more likely to drop out of school permanently with insufficient education. Peru is a poignant example of an economically advancing developing country with such problems in the education sector. Through steady enrollment increases, Peru has almost achieved universal primary school coverage. Nevertheless, educational inefficiency is very high – 18 percent of primary students failed to complete the grade in 2004, and only 73 percent of 12 to 15 year olds had completed the 6-year cycle of primary education in 2003 (MINEDU 2005).

This paper analyses the effect of a reduction in the student-teacher ratio in primary single-teacher schools in rural Peru when a second instructor is added. The effectiveness of teachers should be under close scrutiny since they consume most of the small educational budgets in developing countries. Particularly, researchers still controversially and inconclusively debate about the importance of the student-teacher ratio, i.e., the average number of teachers per student. At the same time, changes in this ratio have huge budgetary consequences and can bind or free up resources for other educational inputs. For example, the World Bank (2007) estimates that in 2005, Peru spent 83 percent of current and 75 percent of total educational expenditure on wages and salaries. Also, it maintains a student-teacher ratio of about 24 which is close to the average of Latin America while its GDP per capita is considerably lower than the Latin American average. The World Bank concludes that Peru's student-teacher ratio may be too low, considering that there is little proof of the effectiveness of more teachers on student outcomes.

Since indicators of enrollment, learning achievement and grade completion are lowest in poor rural communities of developing countries, more teachers may be most effective in such a context, if at all. In sparsely populated rural areas, children often acquire education in small multi-grade schools where teachers teach multiple grades at a time. In the extreme, only one teacher is responsible for the whole school. In these single-teacher schools, the addition of a second teacher reduces multi-grade teaching and class size and may thus be a strong driver of improvement. Theory implies that increases in school inputs have a non-decreasing effect on the level of enrollment and an ambiguous effect on grade completion levels and rates if they improve school quality. Findings on the impact of changes in the student-teacher ratio in schools at the bottom of the quality distribution can inform policymakers on the trade-off between teacher quantity and other educational inputs. While much empirical work has addressed the effect of school inputs on cognitive educational achievement (cf. Hanushek 2003), less effort has been devoted to the equally important questions of their impact on enrollment and school progression.

In order to inform about the effect of lower student-teacher ratios, I employ matched difference-in-difference estimates using a unique longitudinal school census data set from Peru. Difference-in-difference estimation allows understanding the addition of a second teacher as a treatment to single-teacher schools and calculating its effect on educational outcome variables. Before estimating, I employ propensity score matching to mitigate possible bias of results by creating an appropriate control group in observational data.

Matched difference-in-difference estimates show a positive treatment effect on enrollment of about 13 percent which translates into increased grade completion *levels*. The analysis suggests that most of the enrollment effect is caused by lower between-year drop-out *before* introducing the second teacher in treated schools, possibly in anticipation of improved schooling conditions. Via increased enrollment, treated schools produce significantly more grade completers. Nevertheless, the analysis also shows that there is no significant after-treatment effect on grade completion *rates* despite a roughly 40 percent improvement in the mean student-teacher ratio.

There are several possible reasons why the analysis does not show an after-treatment impact on grade completion rates: first, treatment keeps more students in school who would have dropped out in the absence of smaller classes and have high propensities to fail. Second, additional teachers willing to teach in remote areas may be from the bottom of the teacher quality distribution such that teacher training, e.g., on multigrade teaching, may be more effective than reducing class sizes. Indeed, I find that second teachers have relatively more non-permanent positions and work fewer hours even though this may be unrelated to teacher quality. Third, out of school factors rooted in the economic and social environment of children may play a predominant role in poor rural areas, such as low and volatile household incomes. These factors may be unrelated to educational policy and thus harder to address. In summary, increasing teacher quantity is thus unlikely to solve Peru's problem of educational inefficiency.

2.2 Background

2.2.1 Inefficiency in Primary Education in Peru

The Peruvian school system is divided into pre-primary, primary, secondary and higher education. Primary education consists of 6 grades and starts at age 6. In principle, primary and secondary education in Peru is free and compulsory, but households face substantial costs of education[9] and enforcement of attending school is hard in remote areas.

There are three categories of primary school according to the relative number of teachers present: complete, multi-grade, and single-teacher schools. In the first case, the number of teachers equals or exceeds the number of classes. In the second case, at least two teachers are present in school, however, there are more grade levels than teachers thus resulting in grouping of classes. In the last case, there is also multi-grade teaching but only one teacher exists for all students of all grades, typically teaching them altogether in one classroom. Sparsely populated regions, especially in the Andes and the Amazon basin, inhibit appropriate schooling conditions for many students in those remote places. As a result, multi-grade schools are wide-spread (Hargreaves et al. 2001).

Table 2-1 summarizes school characteristics in 2004 by types of public schools, adding private schools as an additional category. Single-teacher and multi-grade schools account for about two thirds of the school universe in Peru and host about one third of students. They are predominantly rural and more than 60 percent of all schools are located in the poorest quintile of districts.[10] Almost all urban schools are complete schools, and more than 90 percent of private schools are urban. All public school types have similar average student-teacher ratios, between 24 and 25, but single-teacher schools have the highest variance: at the 5th percentile there are 10 students per teacher, at the 95th percentile there are 50. Even though the Ministry of Finance has intended an average student-teacher ratio of 35 in urban and 20 in rural areas, with some variations by level and for remote areas[11], both urban and rural schools have close to 25 students per teacher (not shown).

[9] See Saavedra and Suárez (2002).
[10] Poverty was calculated based on the Peruvian census 2005 using a district deficiency index which includes share of illiterate women, children under 12, undernourished people, and households without access to water, electricity, sanitation.
[11] See World Bank (2001).

Table 2-1. Summary Statistics by School Type, 2004

	Public Single-Teacher		Public Multi-Grade		Public Complete		Private (All Types)	
Share of Schools	0.255		0.394		0.167		0.184	
Share of Students	0.068		0.300		0.440		0.192	
Enrollment	24.9	(13.0)	69.5	(45.1)	245.0	(208.3)	95.9	(113.7)
Teachers	1.00	(0.00)	2.89	(1.28)	10.37	(6.48)	7.27	(5.52)
Student-Teacher Ratio	24.90	(12.97)	23.86	(8.72)	24.73	(6.81)	12.48	(8.51)
Lowest Community Quintile	0.66	(0.48)	0.61	(0.49)	0.36	(0.48)	0.05	(0.21)
Rural	0.99	(0.09)	0.96	(0.21)	0.52	(0.50)	0.07	(0.26)
N	8182		12622		5367		5896	

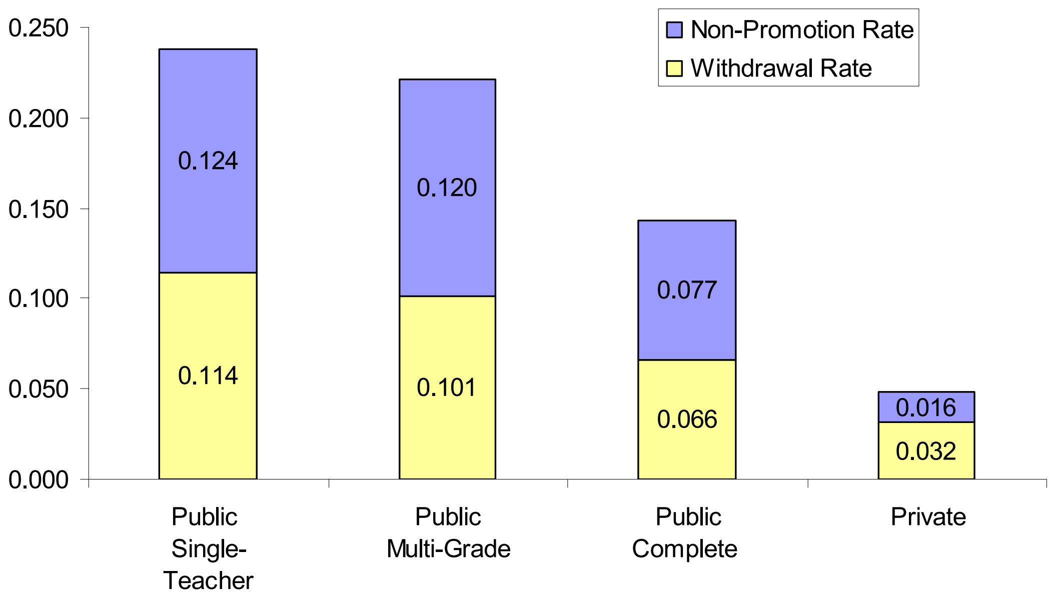
Source: Own estimates based on school census data 2004. Note: Means in the left column, standard deviations in brackets.

Peru has made significant progress in the expansion of primary school coverage for its population. Based on calculations from the national household survey ENAHO, according to the Ministry of Education (MINEDU 2005), in 2003, 96 percent of all children between 6 and 11 years old were enrolled in school. This figure distributes evenly between boys and girls, with a bias towards urban versus rural areas (98 to 93 percent). While among the non-poor, 99 percent of children were enrolled in school, this figure drops to 97 percent for poor children and 93 percent among the extremely poor.

Along with high coverage, educational inefficiency due to drop-out or grade repetition is pervasive in Peru. In 2003 approximately 91 percent of 15- to 17-year-olds had completed primary education, but only 73 percent of 12- to 15-year-olds (MINEDU 2005). Taking into account that primary school can be completed at age 11, many students finish with significant delay. Among the extremely poor, figures are even more drastic, with 54 percent of the population between age 12 and 14, and 78 percent between age 15 and 17, having completed six grades of primary education.

Peru's educational inefficiency resides strongly in high grade non-completion rates (see Figure 2-1). The non-promotion rate refers to the share of students enrolled and showing sufficient attendance but failing the grade due to non-promotion by decision of the teacher. The withdrawal rate denotes the share of students enrolled but failing the grade due to within-year drop-out or insufficient attendance. The sum of non-promotion and withdrawal rate, i.e., the total share of students not completing the grade, is the failure or non-completion rate. In single-teacher and multi-grade schools, more than 20 percent of all students each year fail to complete the grade. Withdrawal and non-promotion contribute almost equally to grade failure. This compares to around 14 percent grade failure in complete multi-grade schools, and 5 percent in private schools. Due to a national average failure rate of 18 percent, schools host many over-aged repeaters and by grade 6, public school students who have not dropped out are on average 1.3 to 1.4 years too old (not shown).

Figure 2-1. Grade Non-Completion Rates by School Type, 2004



Source: Own estimates based on school census data 2004.

There are obvious economic reasons why one should care about drop-out and grade repetition: costs. UNESCO (1998) estimates that in developing countries, between 10 and 40 percent of total public current expenditure on education are spent on wastage before grade 5. Repeaters use more resources such as teaching time, space, textbooks or school meals which may be saved or used for other children, and create a heterogeneity in class that distorts normal instruction. Educational inefficiency puts a burden onto the whole economy, in the form of reduced growth perspectives. In developing countries, this is especially true for rural regions. Drop-out and grade repetition also result in costs at the individual level, e.g., by causing low self-esteem, negative attitudes towards school and higher propensity for criminality. Droppers often relapse into illiteracy. Furthermore, there tends to be a reinforcement of discrimination as children from poorer households often remain uneducated.

2.2.2 School Quality, Inputs, Enrollment and Grade Non-completion

Generally, two causes for school inefficiency can be distinguished: those rooted in economic and social environment of children, *out-of-school* reasons, and those rooted in the school system, *in-school* reasons (Randall, Anderson 1999). This paper concentrates on the latter, specifically teachers as school inputs and their effect on student enrollment and grade completion.

Enrollment

In a simple utility calculation without school choice a student weighs the benefits and costs of completing an additional grade in school (see for example Gertler, Glewwe 1990). Enrollment occurs if the value added from an additional year of schooling is positive, i.e., the benefits of schooling exceed the costs. Benefits of education typically include intrinsic valuation of schooling and the wage return after completing the additional year both of which depend on the quality of education in the respective grade. The costs of an additional year of schooling can be direct, such as school fees, transportation costs, or costs of learning materials, and indirect opportunity costs. A household sending its child to school faces an opportunity cost from losing a worker in the household or labor market. This cost increases in household deprivation and the wage equivalent for the student from not going to school.

Withdrawal

Why would children enroll for school and subsequently withdraw? One possibility is that parameters in the utility calculation change during the year, e.g., with the occurrence of shocks to household wealth, the labor market or school quality. Furthermore, there may be uncertainty about the parameters necessary to decide on the additional year of schooling at the time of the enrollment choice. Students may enroll if their expected value of schooling was positive before enrollment and withdraw during the year if uncertainty is resolved and the resulting utility outcome has turned negative. The quality of schooling may be one of these parameters for which uncertainty resolves after enrollment.[12] If increases in quality trigger increased enrollment, withdrawal rates may rise if the marginal students' propensity to withdraw is higher than that of previous students.

Non-promotion

As for students who do show sufficient attendance for possible promotion to the next grade, probability of promotion should be a non-decreasing function of the student's learning achievement. Learning achievement will weakly increase in educational quality and the student's ability, effort and attendance. Nevertheless, in an attempt to optimize their use of time between leisure, studying and work, students may scale back studying effort as a response to an increase in educational inputs. Also, the decision of being promoted depends on the teacher's assessment of the student's achievement at the end of the grade which may have a non-meritocratic component attached to it.[13] Increases in educational quality thus do not necessarily increase promotion probabilities among students.

A simple theory (cf. Manski 1989) predicts that increases in school quality should increase the returns to schooling and thus have a non-decreasing effect on enrollment. The effect on failure levels and rates is ambiguous because of quality effects on previous students, who would also have enrolled under the old quality level, and newly attracted students: if newly attracted students have some positive probability of failing, failure levels and rates may rise if they are not offset by decreased failure levels and rates of the previous students.[14] As a consequence, increases in quality may increase enrollment and change failure levels and rates in any direction.

School Inputs and School Quality

The quality of an educational system is often measured by its inputs since output is harder to quantify. This neglects the complex process which transforms these educational inputs into outputs. But even by doing so, it is hard to establish a causal relationship

[12] This does not consider the case of a positive probability of failing the school year. Also, there is no distinction of school children according to their ability.
[13] For example, evidence for discrimination based on social background and previous grade repetition has been found for Honduras (Marshall 2003, McGinn et al 1992).
[14] If school quality deteriorates through higher enrollment and this decrease more than offsets the original increase in quality, failure levels may even increase among students who would have enrolled under the old quality level.

between inputs and outputs since outputs may also affect the inputs into the system. For example, UNESCO (1998) correlates average student-teacher ratios in primary education by country with school efficiency as an outcome variable. The coefficient of correlation is -0.65, suggesting a strong influence of the input onto the output. Low school efficiency, however, also influences educational inputs – in this example via the channel of repeating students who clog the system and take up inputs, such as teacher time, away from others.

Although school resources are known to be poor measures of school quality (Hanushek 1995, 2003) the number of teachers is an interesting educational input to study. The student-teacher ratio is a measure of average class size or real resources devoted to schools, and has been used as its proxy in the literature (e.g. Case, Deaton 1999). The amount of teachers per student translates directly into current expenditures on education. Measuring the effect of this crucial input would allow improving their allocation in the context of developing countries with scarce budgetary resources. Teachers are among the most important determinants of children's education and even though teacher quality has been shown to affect outcomes more than teacher quantity (Rivkin et al. 2005), quantity is more easily observable and measurable.

In poor and sparsely populated regions of developing countries, enrollment, class-size and the student-teacher ratio often coincide when the entire student body is taught in one classroom by one teacher. Changes in the number of teachers in this setting can reasonably be assumed to be big changes in school quality: not only does the student-teacher ratio halve with an additional teacher, but students also gain from sharing their teacher with fewer other grades, creating more homogeneous classrooms. In this paper, I thus approximate a change in school quality by the addition of a second teacher to a single-teacher school.

2.2.3 Literature Review

Few analyses on enrollment and grade completion in developing countries examine in-school rather than out-of-school determinants. On the one hand, most of such analyses are performed at the individual level and examine individual, household and community factors, but usually not school characteristics, which drive school progression (e.g., Duryea, et al. 2007 on income risk; Evans, Miguel 2007 on parent death; Pal 2004 on various factors; Meekers, Ahmed 1999 on pregnancy). On the other hand, analyses at the school or class level are often focused on learning achievement, not school progression (e.g., Krishnan et al. 2005 on teacher absence, McEwan 2003 on peer effects; Kingdon 1996 on teacher and school characteristics). For example, international student assessments of learning achievement that sometimes include developing countries, such as PISA, collect tremendous information on individual, classroom and school characteristics at one point in time but not over time for the same observational units. One contrary example of an analysis on school progression using both individual and school-level data over time is Hanushek et al. (2006) who estimate a behavioral model of primary school drop-out behavior. They find that students act on differences in school quality measured as expected achievement improvements, and are more likely to drop out of low quality schools because of relatively lower labor market returns compared to high quality schools.

An additional hindrance to analyses on in-school determinants of enrollment and school progression, such as possibly the student-teacher ratio, is the endogeneity problem due to omitted variable bias and reverse causality. For example, low class-size schools could be high-quality according to many characteristics, of which some are not measured. Also, bureaucrats may react to the output of schools, either by specifically allocating resources to high- or low-efficiency institutions. There are thus only few convincing attempts to estimate the effect of the student-teacher ratio on educational outcomes in developing countries that use particularities of the respective countries' institutions. For example, Case and Deaton (1999) exploit student-teacher ratio differences before the end of Apartheid in South Africa and find strong significant effects on enrollment, attainment and test scores. Angrist and Lavy (1999) exploit discontinuities in class sizes induced by Maimonides' rule in Israel and find significant effects on test scores, but only in some grades. The evidence, however, is far from conclusive.

This paper thus contributes to the aforementioned strand of literature: it provides new evidence on student-teacher ratio effects on indicators of school progression in a developing country using panel data.

2.3 Empirical Implementation

2.3.1 Data

The Peruvian school census is collected on a yearly basis by the statistical unit of the Peruvian Ministry of Education. It covers all Peruvian educational institutions over time with questionnaires specific to the type and level of institution. Information is self-reported to reflect present school registers at the date of May 30. Only information on end-of-year results, such as grade completion, is collected for the previous school year. Thus, one needs to combine the census information of two consecutive years to build a profile of the end-of-year results for students covered in the first year. The analysis uses census information from 2004, 2005 and 2006 to fully cover the years 2004 and 2005 of formal non-adult primary schools.

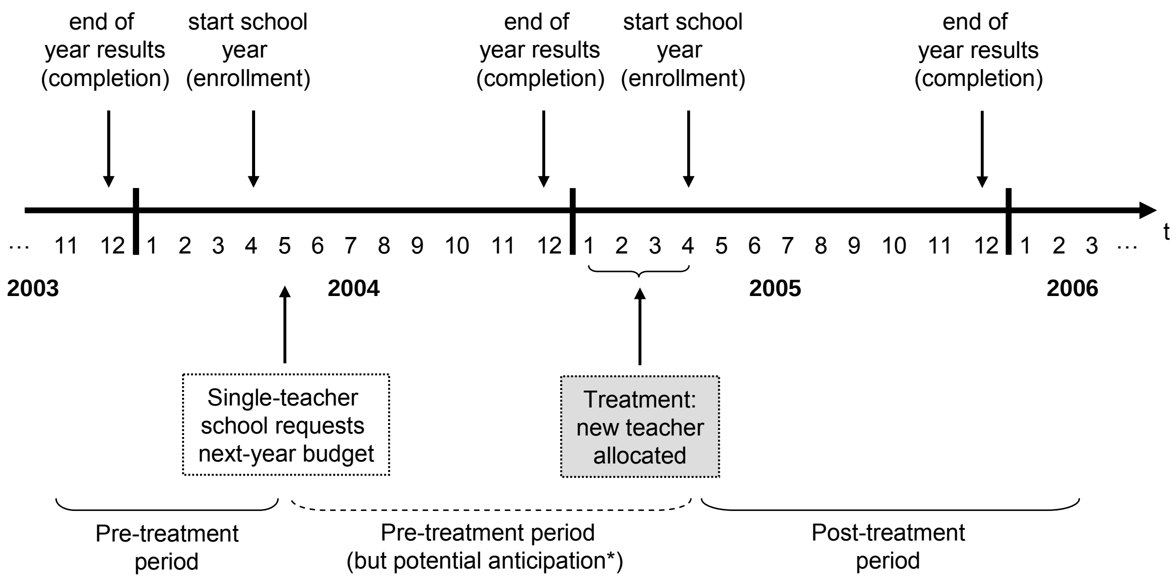
The information does not allow for individual student profiles but aggregation at the grade and school level. For example, information contains the grade structure of students according to gender, age, native language and repeater status but it is not possible to follow who exactly is failing the grade. Teacher information is collected at the school level for primary schools. School infrastructure information is also available but due to a change in questionnaire not comparable between 2004 and 2005.

By use of district identifiers, the school census data is complemented by a data set from 2005 containing district population information and proxy variables for poverty status of the communities, such as the share of households without water access or electricity.

2.3.2 Estimation Strategy and Analytical Framework

Given the difficulties to identify exogenous changes in class size or the student-teacher ratio, I use a quasi-experimental setting outlined in Figure 2-2 focusing on changes in the number of teachers as input changes. These changes, however, may also be prone to result from previous period outcomes, e.g., if additional teachers are allocated to particularly bad schools. This issue is addressed using retrospective data. Although experimental data are often considered more reliable, a retrospective setting does not suffer from a potential "Hawthorne" effect where participants are aware of being in an experiment and thus do not behave naturally (cf. Krueger 1999).

Figure 2-2. Time Line and Treatment Setup



* The possible anticipation of treatment is discussed in the appendix.

First, in order to exploit changes in the number of teachers in Peruvian primary schools, I only consider the sub-sample of schools with a single teacher in the first period, 2004, in rural areas. The schools employ one teacher who is responsible for teaching up to six grades simultaneously with class sizes between a few and several dozen students. This situation is typical for rural poor regions in developing countries which perform worst in enrollment rates, grade completion and learning outcomes and are thus the most interesting unit of analysis. Also, these schools are located in sparsely populated areas characterized by lack of school choice which mitigates concerns of interaction with neighboring schools (Hargreaves et al. 2001, Urquiola 2006).

Second, I consider the addition of a second teacher to rural single-teacher schools as a treatment for which I calculate the average treatment effect on the treated by difference-in-difference estimation. The considered outcomes are enrollment, promotion and failure levels and rates. The reason for analyzing the effect of the second teacher is that an additional teacher promises highest outcome changes in the considered single-teacher schools. Not only is the student-teacher ratio halved, students also enjoy the benefit of sharing their teacher with fewer other grades such that relevant teacher time is more than doubled. The effect of further teachers is likely to be non-linear and decreasing in more teachers. One more teacher in the schools with worst outcomes has thus the least budgetary consequences but the highest possible effect.

The idea of difference-in-difference estimation is to estimate the mean impact of treatment by calculating the difference between changes over time for the treatment and control group. The key assumption concerning selection bias is that the unobserved difference in mean counterfactual outcomes between treated and untreated units is constant over time. If so, outcome changes of the control group disclose the counterfactual outcome changes of the treated units. The assumption may be problematic if treatment units have been specifically selected on the promise of yielding different rates of outcome change than untreated units.

Consequently, we need to understand the process of teacher allocation to schools and the important determinants of this process which may also influence our outcome variables of interest. The budgeting process in the education system is quite fragmented in Peru. Every year in May, one to two months after the beginning of the school year, schools present budget requests for January of the following year to Educational Service Units. They consolidate them for the Regional Directorates, which forward aggregated budgets to Transitional Councils of Regional Administration, which are again consolidated by the Ministry of the Presidency and then presented to the Ministry of Finance (MoF).

For the MoF, the foremost budgetary priority is to cover teacher salaries and pensions before recurrent expenditures may be granted to Regional Directorates for other basic services. The loose formula for allocating teachers to the regions is based on a desired student-teacher ratio of 20 in rural areas and 35 in urban areas. Other educational materials, such as textbooks, are generally bought by the MoF and distributed to Regional Directorates. Afterwards, the Regional Directorates have discretionary power over allocating teachers, budget and materials to the schools before April (World Bank 2001).[15] Alcazar et al. (2002) found that this discretion is used in ways that are non-transparent and cannot necessarily be anticipated by schools, for example by not allocating requested teachers, or allocating teachers beyond what was requested.

If treatment is dependent on first applying for a second teacher and then being allocated sufficient funds by the Regional Directorates, schools which end up with an additional teacher may differ from those which remain single-teacher schools along important dimensions. These dimensions are relevant as long as they influence both treatment allocation and educational outcomes. In order to mitigate the potential bias arising from this selectivity, I employ propensity score matching of single-teacher schools (in 2004) which do and do not receive treatment (in 2005) to construct an appropriate control group along dimensions which may matter both for treatment and outcome. For example, personal connections or distance of the school to the next Regional Directorate may positively influence the probability to receive a second teacher but are probably irrelevant for the success of students. Previous year success of students, however, may influence both, the probability for treatment and outcomes this year. Matching reduces the bias in double-difference estimates by eliminating initial heterogeneity of observables between the treatment and comparison group. The method is superior to propensity score matching which assumes conditional exogeneity of unobservables with respect to treatment status conditional on observables and is prone to suffer from selection bias based on latent variables (Ravallion 2007).

On the matched and pooled sample, I estimate the difference-in-difference OLS equation (1) where the null hypothesis states that treatment does not have an effect on outcome.[16] Outcomes can be the level of enrollment, the level of grade completers and failers as well as the share of completers and failers. The equation is of the form

$$(1) \qquad Y_{st} = \beta_0 + \beta_1 T_s + \beta_2 P_t + \beta_3 T_s * P_t + \beta_4 X_{st} + e_{st}$$

where the outcome (Y) in school (s) and year (t) is a function of being in the teacher-treatment group (T), a post-treatment dummy for the year 2005 (P), the

[15] As school registers are reported at May 30, they incorporate the number of teachers for the whole year; even though I cannot exclude with certainty that no more teachers are added after May, this is highly implausible.
[16] As a robustness check, I estimate the difference-in-difference equation by tobit since the dependent variable is censored between 0 and 1 (see appendix).

interaction effect between being in the treatment group and being in the second year (T * P), a vector of control variables (X) which are mostly also used for matching, and a random error term (e). β_3 is the main coefficient of interest, the *average treatment effect on the treated* (ATT).

2.3.3 Propensity Score Matching

Table 2-2 summarizes the variables used for propensity score matching between the single-teacher schools in 2004 that do and do not receive an additional teacher in 2005, i.e., the raw treatment and control group. On average, treated schools are significantly larger and thus more likely to receive an additional teacher. A t-test of mean comparison also reveals that treatment and control group differ along other dimensions at the 10 percent significance level, such as teacher gender, withdrawal share, age heterogeneity of students, existence of a parent committee, and the district share of households with sanitation.

Table 2-2. Summary Statistics – Unmatched Treatment and Control Groups before Treatment (2004)

	Treatment Group		Control Group		T-Test p-value
Enrollment	39.095	(15.36)	23.861	(11.15)	0.000
Teacher w/ Permanent Contract	0.774	(0.42)	0.753	(0.43)	0.380
Teacher Male	0.580	(0.49)	0.522	(0.50)	0.031
Teacher Hours	38.011	(4.00)	38.343	(3.78)	0.106
Teacher w/ Teaching Degree	0.815	(0.39)	0.816	(0.39)	0.953
Non-Promotion Share (lagged)	0.129	(0.10)	0.125	(0.11)	0.497
Withdrawal Share (lagged)	0.127	(0.12)	0.126	(0.12)	0.935
Non-Promotion Share	0.123	(0.10)	0.120	(0.11)	0.692
Withdrawal Share	0.105	(0.10)	0.117	(0.12)	0.059
Share Repeaters in Class	0.051	(0.08)	0.055	(0.09)	0.420
Share Reentrants in Class	0.123	(0.11)	0.121	(0.12)	0.794
S.D. Age Distribution	2.421	(0.48)	2.289	(0.53)	0.000
Share Working	0.137	(0.34)	0.127	(0.32)	0.598
Share Not First Language	0.164	(0.35)	0.185	(0.36)	0.269
Share Male	0.508	(0.08)	0.516	(0.12)	0.263
Morning Classes	0.793	(0.41)	0.816	(0.39)	0.277
Food Program	0.736	(0.44)	0.697	(0.46)	0.114
Health Service	0.144	(0.35)	0.153	(0.36)	0.658
Language Other Native	0.071	(0.26)	0.062	(0.24)	0.486
Language Quechua	0.057	(0.23)	0.060	(0.24)	0.817
Bilingual School	0.128	(0.34)	0.143	(0.35)	0.441
Parents Committee	0.173	(0.38)	0.212	(0.41)	0.077
Rural	1.000	(0.00)	1.000	(0.00)	.
Share No Water (D)	0.557	(0.25)	0.541	(0.25)	0.251
Share No Sanitation (D)	0.411	(0.23)	0.455	(0.24)	0.001
Share No Electricity (D)	0.671	(0.24)	0.661	(0.25)	0.451
Share Illiterate Women (D)	0.252	(0.14)	0.262	(0.14)	0.208
Share Children 0-12 (D)	0.332	(0.05)	0.333	(0.05)	0.686
Share Malnutrition '99 (D)	0.439	(0.13)	0.440	(0.13)	0.980
N	367		5183		

Source: Own estimates based on school census data 2004 only for schools with full set of control variables available. Note: Means in the left column, standard deviations in brackets. (D) denotes variables measured at the district level.

Table 2-3 shows the results of a probit analysis of treatment on the vector of observed control variables. In line with the MoF budgeting rules, enrollment has the biggest influence on the probability of being allocated a second teacher. Additionally, the following characteristics are significantly correlated with receiving a second teacher: having a lower withdrawal share, having a lower share of repeaters, having a higher age heterogeneity of students, offering school meals, being a native Quechua school, and being located in a district with higher rates of illiteracy among women and higher shares of children.

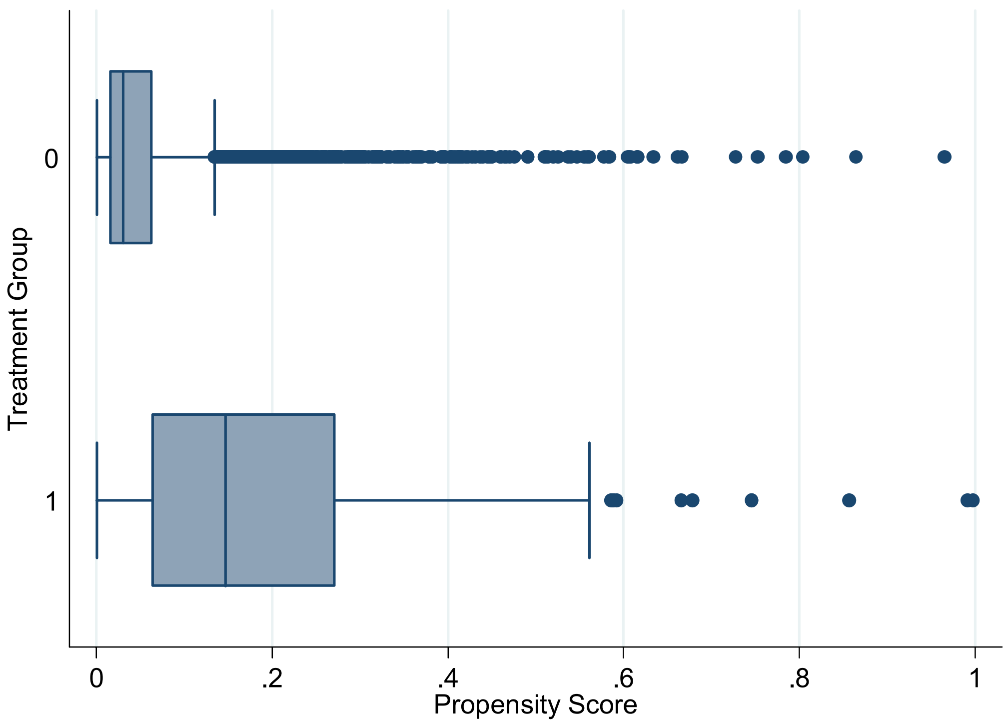
Table 2-3. Probit Regression of School Treatment

	Coefficient		S.E.
Enrollment	0.048	***	(0.002)
Teacher w/ Permanent Contract	0.119		(0.076)
Teacher Male	0.069		(0.063)
Teacher Hours	-0.001		(0.008)
Teacher w/ Teaching Degree	0.051		(0.086)
Non-Promotion Share (lagged)	-0.096		(0.427)
Withdrawal Share (lagged)	0.433		(0.326)
Non-Promotion Share	-0.484		(0.316)
Withdrawal Share	-1.173	***	(0.327)
Share Repeaters in Class	-0.909	**	(0.459)
Share Reentrants in Class	-0.097		(0.403)
S.D. Age Distribution	0.185	***	(0.063)
Share Working	0.005		(0.094)
Share Not First Language	0.126		(0.094)
Share Male	-0.479		(0.301)
Morning Classes	0.007		(0.083)
Food Program	0.131	*	(0.073)
Health Service	-0.020		(0.086)
Language Other Native	0.091		(0.148)
Language Quechua	0.333	**	(0.144)
Bilingual School	-0.045		(0.109)
Parents Committee	-0.076		(0.080)
Share No Water (D)	0.051		(0.155)
Share No Sanitation (D)	0.076		(0.158)
Share No Electricity (D)	0.194		(0.170)
Share Illiterate Women (D)	-1.047	***	(0.364)
Share Children 0-12 (D)	-5.027	***	(0.975)
Share Malnutrition '99 (D)	0.214		(0.392)
Constant	-1.535	***	(0.450)
R-Squared	0.198		
Observations	5309		

Source: Own estimates based on school census data 2004. Note: Robust standard errors in brackets, significance levels: * p<0.10, ** p<0.05, *** p<0.01

If additional teachers are placed where outcomes are higher to begin with, or where the district has higher shares of literate women and lower fertility, there might be bias in difference-in-difference estimates, arising from the fact that the targeted schools may be able to show higher rates of productivity growth than their peers. As far as observable characteristics are concerned, this worry is taken care of by finding an appropriate control group via propensity score matching and making sure that covariates are balanced between the groups. A remaining concern to identification is heterogeneity between treatment and control group with respect to unobservables that induce different rates of outcome growth over time, conditional on treatment.

Figure 2-3. Distribution of Propensity Score by Treatment Status



Source: Own estimates based on school census data 2004.

Figure 2-3 shows the distribution of the estimated propensity score by treatment status. Non-treated schools are concentrated heavily in the lowest score quintile with a median of 0.03. However, due to the large number of non-treated schools, close matches for treated schools can even be found in high-score regions. Table 2-4 shows the resulting treatment and control group after performing nearest neighbor matching with 5 neighbors and replacement. The treatment group reduces to 348 schools after eliminating observations off the common support and those without an appropriate neighbor within a caliper of 0.01. The control group consists of 1071 schools using some more than once as an appropriate match. T-tests can never reject equality of weighted means between treatment and control group variables at the 10 percent significance level.[17] For an elaboration of the characteristics of new compared to old teachers, see the appendix.

Table 2-4. Summary Statistics – Matched Treatment and Control Groups before Treatment (2004)

	Trimmed Treatm. Group		Matched Control Group		T-Test p-value
Enrollment	38.198	(13.90)	38.622	(15.06)	0.657
Teacher w/ Permanent Contract	0.787	(0.41)	0.765	(0.42)	0.419
Teacher Male	0.583	(0.49)	0.568	(0.50)	0.627
Teacher Hours	38.046	(3.97)	38.072	(3.95)	0.921
Teacher w/ Teaching Degree	0.816	(0.39)	0.825	(0.38)	0.723
Non-Promotion Share (lagged)	0.129	(0.10)	0.133	(0.11)	0.590
Withdrawal Share (lagged)	0.127	(0.12)	0.134	(0.11)	0.381
Non-Promotion Share	0.123	(0.10)	0.123	(0.10)	0.979
Withdrawal Share	0.104	(0.10)	0.108	(0.10)	0.523
Share Repeaters in Class	0.051	(0.08)	0.055	(0.08)	0.493
Share Reentrants in Class	0.126	(0.11)	0.130	(0.11)	0.537
S.D. Age Distribution	2.429	(0.48)	2.444	(0.47)	0.639
Share Working	0.138	(0.34)	0.157	(0.35)	0.411
Share Not First Language	0.161	(0.35)	0.176	(0.36)	0.526
Share Male	0.507	(0.08)	0.506	(0.09)	0.790
Morning Classes	0.793	(0.41)	0.785	(0.41)	0.749
Food Program	0.753	(0.43)	0.766	(0.42)	0.640
Health Service	0.149	(0.36)	0.167	(0.37)	0.465
Language Other Native	0.069	(0.25)	0.081	(0.27)	0.484
Language Quechua	0.060	(0.24)	0.064	(0.25)	0.823
Bilingual School	0.129	(0.34)	0.148	(0.36)	0.425
Parents Committee	0.178	(0.38)	0.173	(0.38)	0.851
Rural	1.000	(0.00)	1.000	(0.00)	.
Share No Water (D)	0.554	(0.25)	0.556	(0.25)	0.899
Share No Sanitation (D)	0.412	(0.23)	0.417	(0.25)	0.736
Share No Electricity (D)	0.673	(0.24)	0.676	(0.25)	0.853
Share Illiterate Women (D)	0.254	(0.14)	0.263	(0.14)	0.309
Share Children 0-12 (D)	0.332	(0.05)	0.335	(0.05)	0.425
Share Malnutrition '99 (D)	0.441	(0.13)	0.447	(0.13)	0.493
N	348		1071		
Sum of Weights	348		348		

Source: Own estimates based on school census data 2005. Note: Means in the left column, standard deviations in brackets. (D) denotes variables measured at the district level.

[17] Estimation results in the next section are qualitatively very similar when using nearest-neighbor matching with 1 neighbor or kernel matching. Results are available from author upon request.

One could argue that end-of-year results from 2004 should not be included in the matching procedure since they may be potentially influenced by the anticipation of receiving a second teacher, given that the budgeting process for the 2005 starts mid-2004. An extended discussion of this issue and further estimation results are presented in the appendix.

2.4 Results

Table 2-5 to Table 2-9 show the results of estimating equation (1), i.e., difference-in-difference estimates with robust standard errors on the pooled sample of schools on different measures of outcome: enrollment, the student-teacher ratio, completion, failure, non-promotion and withdrawal levels and rates at the school level.[18]

Table 2-5. Matched Difference-in-Difference Estimates: Treatment Effect on Enrollment and the Student-teacher Ratio

	[1]	[2]	[3]	[4]
Dependent Variable	Enrollment		Student-teacher ratio	
Treatment*2005	4.728***	4.780***	-15.635***	-15.592***
	(1.419)	(1.284)	(1.221)	(1.111)
Treatment Group	-0.424	-0.008	-0.424	-0.042
	(0.955)	(0.860)	(0.955)	(0.860)
Year 2005	-2.200**	-2.193***	-2.200**	-2.182***
	(0.873)	(0.776)	(0.873)	(0.779)
Controls	No	Yes	No	Yes
Adj.R-Squared	0.009	0.217	0.241	0.399
Observations	2838	2838	2838	2838

Source: Own estimates based on school census data. Note: Robust standard errors in brackets, significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 2-5 shows an important finding: the addition of a second teacher to rural single-teacher schools increases enrollment significantly: on average by 4.7 students (column 1). At 38.2 students before treatment this represents an enrollment increase of

[18] Control variables are used as follows: regressions on enrollment and outcome *levels* include school and district variables; regressions on outcome *shares* include student, school and district variables. Student level: standard deviation of age distribution, share working, share with other first language than school language; School level binary variables: morning classes, food program, health service, language Quechua, language other native, bilingual school, parents committee; District level: share households without water, without sanitation, without electricity, share of illiterate women, share of children age 0-12, share of malnourished children in 1999.

about 13 percent. The effect is almost the same (4.8 students) when including the set of control variables (column 2). At the same time, the student-teacher ratio drops considerably – the treatment effect of -15.6 students per teacher (columns 3 and 4) is equivalent to a decrease of roughly 40 percent.

There are several possibilities why enrollment increases in treated schools: (i) schools are allocated a second teacher when there is a large cohort one year before enrollment in first grade, (ii) treated schools attract students from other schools, (iii) treated schools attract formerly not enrolled students, and (iv) treated schools have lower drop-out rates between 2004 and 2005.

(i) Table 2-6 shows that treatment is not allocated in anticipation of a large new cohort in grade 1. The dependent variable is enrollment in grades 1 to 6 (columns 1 to 6). If (i) was the case we would observe most of the treatment effect in grade 1. Instead, we observe that the enrollment effect is spread out over 5 of 6 grades.

Table 2-6. Matched Difference-in-Difference Estimates: Treatment Effect on Enrollment in Grades 1 to 6

Enrollment in grades 1-6	[1]	[2]	[3]	[4]	[5]	[6]
Treatment*2005	1.194***	1.012**	0.134	0.785***	0.769**	0.886***
	(0.421)	(0.417)	(0.342)	(0.303)	(0.303)	(0.256)
Treatment Group	-0.082	-0.391	0.212	0.132	0.279	-0.158
	(0.297)	(0.283)	(0.241)	(0.208)	(0.210)	(0.175)
Year 2005	-1.212***	-0.448	-0.286	-0.096	-0.101	-0.050
	(0.270)	(0.277)	(0.192)	(0.178)	(0.180)	(0.156)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj.R-Squared	0.122	0.183	0.165	0.116	0.089	0.090
Observations	2838	2838	2838	2838	2838	2838

Source: Own estimates based on school census data. Note: Robust standard errors in brackets, significance levels: * p<0.10, ** p<0.05, *** p<0.01.

(ii) Table 2-7 indicates that the enrollment effect is not merely due to attracting students from other schools. The dependent variable is enrollment, split up into different student categories: students that were promoted from a lower grade or enter grade 1 for the first time (column 1), students that repeat last year's grade due to non-promotion (column 2) or withdrawal (column 3) or that were reincorporated after not being enrolled (column 4) the year before. Panel A shows students coming from a different school, panel

B those coming from the same school. The sum of effects in panel A shows that treatment attracts only about 0.7 students from other schools on average. The bulk of increasingly enrolled students is thus not just pulled away from other schools. This is reasonable as the sampled rural schools can be considered monopolistic entities, far away from other schools.

Table 2-7. Matched Difference-in-Difference Estimates: Enrollment Effects by Student Status

Panel A: Students from Different School

Origin of student: different school	[1] Promoted	[2] Repeater	[3] Reentrant	[4] Reincorpo-rated
Treatment*2005	0.535*	0.048	0.071	0.029
	(0.279)	(0.094)	(0.100)	(0.047)
Treatment Group	0.150	0.036	0.036	0.024
	(0.193)	(0.074)	(0.051)	(0.036)
Year 2005	-0.289**	-0.103*	-0.077**	-0.036**
	(0.125)	(0.053)	(0.035)	(0.016)
Controls	Yes	Yes	Yes	Yes
Adj.R-Squared	0.030	0.013	0.006	0.005
Observations	2838	2838	2838	2838

Panel B: Students from Same School

Origin of student: same school	[1] Promoted	[2] Repeater	[3] Reentrant	[4] Reincorpo-rated
Treatment*2005	2.457**	0.977**	0.510*	0.095
	(1.076)	(0.420)	(0.300)	(0.084)
Treatment Group	0.592	-0.189	-0.299	-0.080
	(0.728)	(0.301)	(0.227)	(0.072)
Year 2005	0.110	-0.785***	-0.392*	-0.157***
	(0.656)	(0.255)	(0.201)	(0.057)
Controls	Yes	Yes	Yes	Yes
Adj.R-Squared	0.174	0.140	0.050	0.011
Observations	2838	2838	2838	2838

Source: Own estimates based on school census data. Note: Robust standard errors in brackets, significance levels: * p<0.10, ** p<0.05, *** p<0.01. The category "promoted within same school" also contains first graders who enter for the first time.

(iii) Also, the treatment effect does not work by attracting students that were previously not enrolled. The effects of column 4 in panel A and B are close to zero and insignificant.

(iv) Instead, most of the increased enrollment effect is found in columns 1 to 3 of panel B, i.e., enrollment is increased by students of the treated school that were promoted in the year before (2.5) or that repeat the grade due to non-promotion (1.0) or withdrawal (0.5). Increased enrollment in treated school thus results from fewer students dropping out between 2004 and 2005 compared to untreated schools, possibly in anticipation of increased school quality. Note that some of these students would have dropped out without treatment anticipation even though they had completed the previous grade.

Table 2-8. Matched Difference-in-Difference Estimates: Completion and Failure Levels

	[1]	[2]	[3]	[4]
Dependent Variable	Completed	Failed	Non-Promoted	Withdrawn
Treatment*2005	4.100***	0.680	-0.136	0.798*
	(1.050)	(0.629)	(0.421)	(0.413)
Treatment	0.124	-0.131	0.191	-0.313
	(0.713)	(0.448)	(0.305)	(0.289)
Year 2005	-2.013***	-0.180	0.088	-0.263
	(0.643)	(0.400)	(0.264)	(0.252)
Controls	Yes	Yes	Yes	Yes
Adj.R-Squared	0.159	0.154	0.194	0.065
Observations	2838	2838	2838	2838

Source: Own estimates based on school census data. Note: Robust standard errors in brackets, significance levels: * p<0.10, ** p<0.05, *** p<0.01

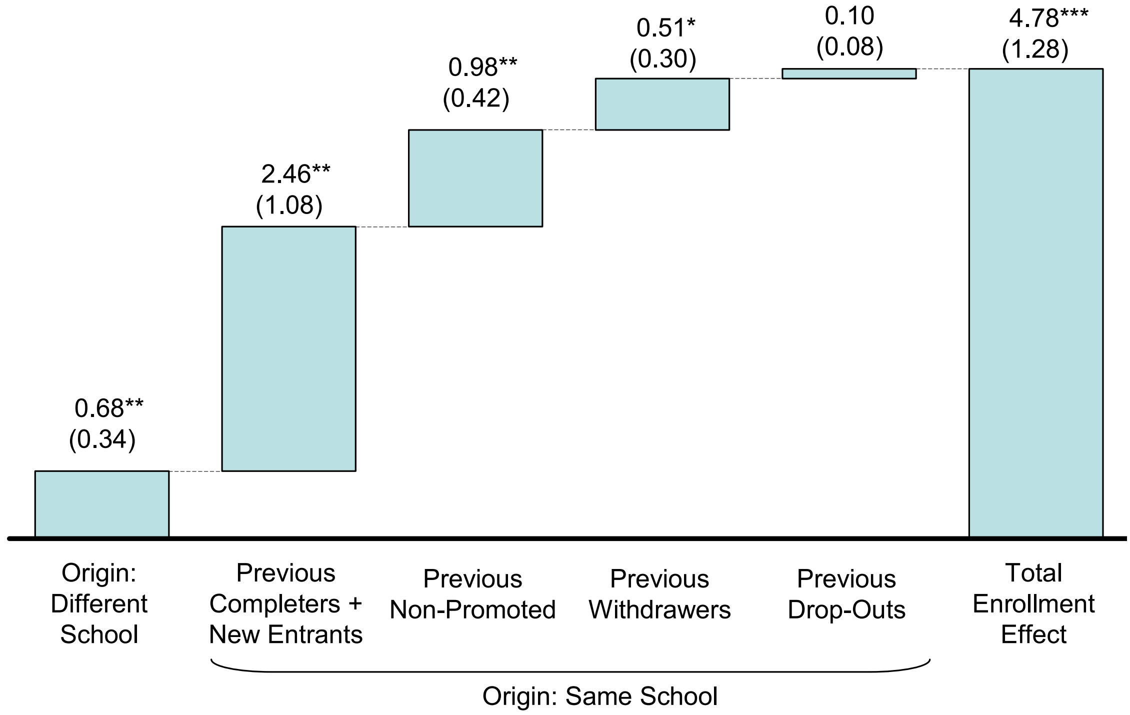
Table 2-8 shows the results for *levels* of grade completion (column 1), failure (column 2), non-promotion (column 3) and withdrawal (column 4) with control variables. Treatment effects from column 1 and 2 add up to the enrollment effect of column 2, Table 2-5.[19] Increased enrollment levels due to treatment thus translate into increased grade completion and failure levels. Column 1 shows that the treatment interaction is positive and significant, i.e., there is an estimated effect of 4.1 additional grade completers after treatment according to the specification including control variables. Similarly, there are an estimated 0.7 additional grade failers after treatment who seem to have a higher tendency for withdrawal than non-promotion. Figure 2-4 decomposes the enrollment

[19] The treatment effects on non-promoted and withdrawn students (columns 3 and 4) do not exactly add up to the treatment effect on failed students (column 2) due to a small number of deceased students who fail the grade but are neither considered non-promoted nor withdrawers.

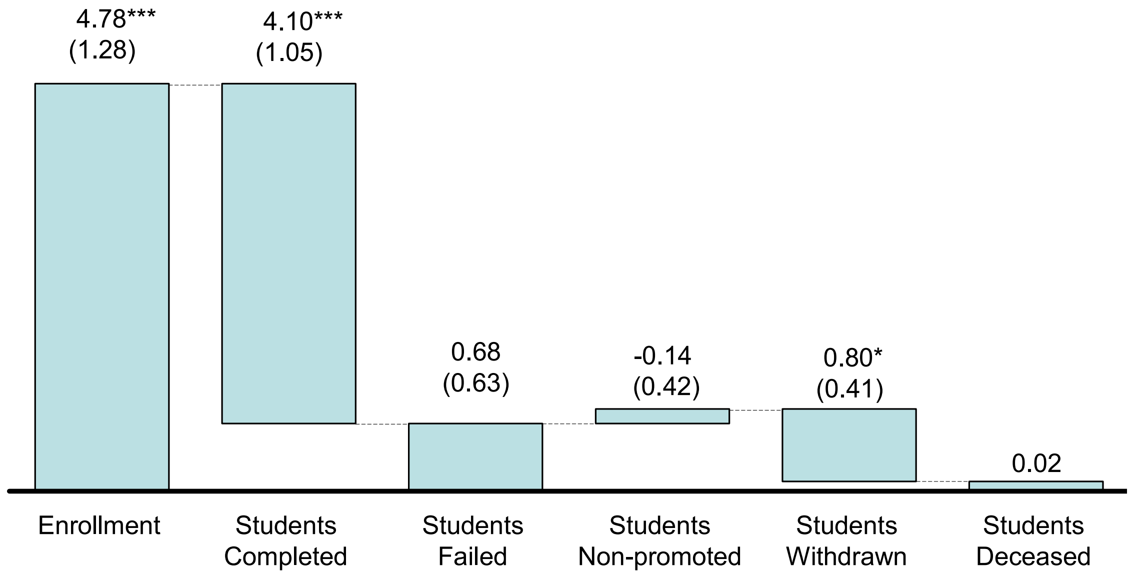
effect graphically; panel A displays the origin of 'additional' students, panel B displays the end-of-year results resulting from the treatment effect on enrollment.

Figure 2-4. Decomposition of Enrollment Effect

Panel A: Origin of Additional Students



Panel B: End-of-Year Results of Additional Students



Source: Own estimates based on school census data. Note: Robust standard errors in brackets, significance levels: * p<0.10, ** p<0.05, *** p<0.01. The category "promoted within same school" also contains first graders who enter for the first time.

Table 2-9 documents the effects of treatment on grade completion and failure *shares*. Contrary to the treatment interaction effect on completion levels, the effect on the completion share is insignificant. Point estimates indicate that the effect of receiving a second teacher is a 1.8 percentage point increase in the completion share. Given standard errors, we can reject with 95 percent confidence that the improvement of the completion rate due to treatment is larger than about 4.2 percentage points or, in other words, less than a fifth of the effect needed to close the gap towards 100 percent completion rate.

Table 2-9. Matched Difference-in-Difference Estimates: Completion and Failure Shares

	[1]	[2]	[3]	[4]
Dependent Variable	Share Completed	Share Failed	Share Non-Promoted	Share Withdrawn
Treatment*2005	0.018	-0.018	-0.015*	-0.003
	(0.012)	(0.012)	(0.009)	(0.009)
Treatment	0.001	-0.001	0.003	-0.003
	(0.009)	(0.009)	(0.006)	(0.006)
Year 2005	-0.015**	0.015**	0.008	0.007
	(0.007)	(0.007)	(0.005)	(0.005)
Controls	Yes	Yes	Yes	Yes
Adj.R-Squared	0.092	0.092	0.126	0.055
Observations	2838	2838	2838	2838

Source: Own estimates based on school census data. Note: Robust standard errors in brackets, significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Columns 3 and 4 of Table 2-8 and Table 2-9 indicate that there may have been a small shift from non-promotion to withdrawal due to treatment. This could be support for the hypothesis that there may be uncertainty about a possible change in the quality of education due to a second teacher. This uncertainty may induce children to enroll and subsequently drop out if the quality is insufficient to provide positive utility.

Note that the treatment effect is most probably not a strictly causal effect from reduced class sizes due to a second teacher. After treatment, there are two presumably counteracting effects: First, we presume a positive effect through class size reduction. Second, there may be a negative side effect in that more students are in school who would not have enrolled in absence of treatment and who may thus be more likely to fail than their peers given lower ability or utility from schooling. Since this latter variable cannot be observed, it may partly be reflected in the treatment effect. The net impact of these two effects shows up insignificantly in the after-treatment effect on completion shares.[20] Nevertheless, this is precisely the effect we are interested in. If the policy objective is to close the gap towards universal primary completion, exactly those children with lowest utility from schooling need to be drawn into school and incentivized to stay there. We thus want to know if this can be achieved by class reductions, and the results do not support this conjecture.

2.5 Conclusion

A matched difference-in-difference analysis presented in this paper shows that the addition of a second teacher to rural primary single-teacher schools in Peru increases enrollment by about 14 percent mainly because of fewer between-year drop-outs in treated schools, possibly in anticipation of higher future school quality. Consequently, completion *levels* in the second period are significantly increased after treatment due to this enrollment effect.

Nevertheless, the analysis also shows that there is no additional significant after-treatment effect on completion *rates*. This disappointing result is a net effect of decreasing the student-teacher ratio on average by about 40 percent and inducing more students to enroll who would not have remained in school in the absence of smaller classes and are thus the first ones to fail. Unfortunately, we cannot say how strong these effects are respectively.

Since rural single-teacher schools tend to lay in the poorest areas improvements in school quality decrease drop-out and thus increase enrollment among those groups of the population where it is lowest. In-school factors thus seem to matter for those parts of the population that are so far excluded from the educational system. This finding is partly along the lines of Hanushek et al. (2006): increased school quality keeps children in school longer; even though in this case input quantity is raised, not output quality. Nevertheless, the finding implies that increases in quality would have to be even more significant and costly to close the gap towards universal primary education. Even though this could not

[20] For this reason, it is not possible to use the treatment interaction as an instrument for the student-teacher ratio to estimate an actual *student-teacher ratio effect* (instead of a treatment effect) using two stage least squares.

be explicitly tested in the paper, it may be a more promising avenue to invest in teacher quality rather than quantity, such as training specifically designed to deal with multi-grade teaching.

Furthermore, the results also suggest that out-of-school reasons may be significant determinants of grade non-completion rates in poor rural areas in developing countries. Households with financial constraints or volatile income flows face high opportunity costs of sending their children to school permanently. These constraints may be lifted by programs such as conditional cash transfers, which support the poorest families financially conditional on sending their children to school. Estimating the impact of such measures on educational inefficiency remains an imperative area for future research.

References

Alcazar, Lorena, Wachtenheim, Erik, López-Cálix, José R. (2002). Public Expenditure Tracking Survey - The Education Sector in Peru. Lima: Instituto Apoyo. Retrieved July 27, 2008, from http://siteresources.worldbank.org/INTPUBSERV/Resources/peru.PETS.eduction.report.sep.2002.pdf.

Angrist, Joshua D., Lavy, Victor (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics, 114*(2), pp. 533–575.

Case, Anne, Deaton, Angus (1999). School Inputs and Educational Outcomes in South Africa. *Quarterly Journal of Economics, 114*(3), pp. 1047–1084.

Duryea, Suzanne, Lam, David, Levison, Deborah (2007). Effects of Economic Shocks on Children's Employment and Schooling in Brazil. *Journal of Development Economics, 84*(1), pp. 188–214.

Evans, David K., Miguel, Edward (2007). Orphans and Schooling in Africa: A Longitudinal Analysis. *Demography, 44*(1), pp. 35–57.

Gertler, Paul, Glewwe, Paul (1990). The Willingness to Pay for Education in Developing Countries: Evidence from Rural Peru. *Journal of Public Economics, 42*(3), pp. 251–275.

Hanushek, Eric A. (1995). Interpreting Recent Research on Schooling in Developing Countries. *World Bank Research Observer, 10*(2), pp. 227–246.

Hanushek, Eric A. (2003). The Failure of Input-Based Schooling Policies. *Economic Journal, 113*(485), pp. F64-98.

Hanushek, Eric A., Lavy, Victor, Hitomi, Kohtaro (2006). Do Students Care about School Quality? Determinants of Dropout Behavior in Developing Countries. National Bureau of Economic Research Working Paper Series No. 12737.

Hargreaves, Eleanore et al. (2001). Multigrade Teaching in Peru, Sri Lanka and Vietnam: an Overview. *International Journal of Educational Development, 21*(6), pp. 499–520.

Kingdon, Geeta G. (1996). The Quality and Efficiency of Public and Private Schools: A Case Study of Urban India. *Oxford Bulletin of Economics and Statistics, 58*(1), pp. 55–80.

Krishnan, Pramila, Habyarimana, James, Dercon, Stefan, Das, Jishnu (2005). Teacher Shocks and Student Learning: Evidence from Zambia. World Bank Policy Research Working Paper Series No. 3602. World Bank.

Krueger, Alan B. (1999). Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics, 114*(2), pp. 497–532.

Manski, Charles F. (1989). Schooling as Experimentation: A Reappraisal of the Postsecondary Dropout Phenomenon. *Economics of Education Review, 8*(4), pp. 305–312.

Marshall, Jeffery H. (2003). Grade Repetition in Honduran Primary Schools. *International Journal of Educational Development, 23*(6), pp. 591–605.

McEwan, Patrick J. (2003). Peer Effects on Student Achievement: Evidence from Chile. *Economics of Education Review, 22*(2), pp. 131–141.

McGinn, Noel et al (1992). Why Do Children Repeat Grades? A Study of Rural Primary Schools in Honduras. BRIDGES Research Report Series No. 13.

Meekers, Dominique, Ahmed, Ghyasuddin (1999). Pregnancy-Related School Dropouts in Botswana. *Population Studies, 53*(2), pp. 195–209.

MINEDU (2005). *Indicadores de la Educación. Perú 2004*. Lima: Ministerio de la Educación.

Pal, Sarmistha (2004). Child Schooling in Peru: Evidence from a Sequential Analysis of School Progression. *Journal of Population Economics, 17*(4), pp. 657–680.

Randall, Laura, Anderson, Joan B. (1999). *Schooling for Success: Preventing Repetition and Dropout in Latin American Primary Schools.* Armonk, NY: Sharpe.

Ravallion, Martin (2007). Chapter 59 Evaluating Anti-Poverty Programs. In T. Paul Schultz and John A. Strauss (Ed.), *Handbook of Development Economics.* Volume 4, pp. 3787–3846. Elsevier.

Rivkin, Steven G., Hanushek, Eric A., Kain, John F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica, 73*(2), pp. 417–458.

Saavedra, Jaime, Suárez Pable (2002). El financiamiento de la educación pública en el Perú: el rol de las familias. Documento de trabajo No. 38. Lima: GRADE.

UNESCO (1998). *Wasted Opportunities: When Schools Fail: Repetition and Drop-out in Primary Schools.* Paris: EFA Forum Secretariat; UNESCO.

Urquiola, Miguel (2006). Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia. *Review of Economics and Statistics, 88*(1), pp. 171–177.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Mass.: MIT Press.

World Bank (2001). *Peruvian Education at a Crossroads: Challenges and Opportunities for the 21st Century.* Washington, DC: World Bank.

World Bank (2007). *Toward High-quality Education in Peru: Standards, Accountability, and Capacity Building.* Washington, DC: World Bank.

Appendix 2

Appendix 2.1 Comparison of New and Old Teacher Characteristics

Could it be possible that the new teachers are different from old teachers, for example because only teachers of the lowest quality go to remote rural areas? Table 2-10 addresses this issue by comparing average observable teacher characteristics – gender, contract and education – between treatment and control group before and after treatment. The upper panel A shows that characteristics before treatment are balanced due to trimming of the treatment group and selection of appropriate control schools via propensity score matching.

Table 2-10. Average Teacher Characteristics Before and After Treatment

	Trimmed Treatm. Group		Matched Control Group		T-Test p-value
A. Before Treatment					
Teacher Male	0.583	(0.49)	0.568	(0.50)	0.627
Teacher Female	0.417	(0.49)	0.432	(0.50)	0.627
Teacher with Teaching Degree	0.816	(0.39)	0.825	(0.38)	0.723
Teacher without Teaching Degree	0.184	(0.39)	0.175	(0.38)	0.723
Teacher with Permanent Contract	0.787	(0.41)	0.765	(0.42)	0.419
Teacher with Fixed Term	0.213	(0.41)	0.235	(0.42)	0.419
Teacher Hours	38.046	(3.97)	38.072	(3.95)	0.921
B. After Treatment					
Teacher Male	0.500	(0.39)	0.554	(0.50)	0.054
Teacher Female	0.500	(0.39)	0.446	(0.50)	0.054
Teacher with Teaching Degree	0.839	(0.30)	0.844	(0.36)	0.795
Teacher without Teaching Degree	0.161	(0.30)	0.156	(0.36)	0.795
Teacher with Permanent Contract	0.632	(0.36)	0.729	(0.45)	0.000
Teacher with Fixed Term	0.368	(0.36)	0.271	(0.45)	0.000
Teacher Hours	34.155	(3.12)	38.115	(4.24)	0.000

Source: Own estimates based on school census data. Note: Robust standard errors in brackets.

About 58 percent of teachers are male, more than 80 percent have obtained a teaching degree and almost 80 percent of teachers have a permanent contract. Also, most teachers work the maximum of 40 hours. In contrast, panel B suggests that second teachers are more predominantly female and work on fixed rather than permanent contracts. Also, additional teachers work fewer hours than their colleagues such that the treatment effect on the number of *students per teacher hour* is less the effect on the student-teacher ratio.[21] While teacher education does not significantly differ after treatment between treated and control schools, fixed term contracts may have a different motivational effect than permanent appointments. It is unknown if these differences indicate lower teacher quality of second teachers which could explain the lack of clear positive effects on grade completion rates.

Appendix 2.2 Tobit Estimates

Since completion and failure shares are censored between 0 and 1, OLS estimation is similar to specifying a linear probability model, including its well-known shortcomings, such as predictions below 0 or above 1 (cf. Wooldridge 2002). As a robustness check, I thus estimate a tobit model with censoring at 0 and 1. The results in Table 2-11 are very close to OLS results (Table 2-9) and indicate that OLS results seem to be an appropriate approximation for the problem at hand.

Table 2-11. Matched Difference-in-Difference Tobit Estimates: Completion and Failure Shares

Dependent Variable	[1] Share Completed	[2] Share Failed	[3] Share Non-Promoted	[4] Share Withdrawn
Treatment*2005	0.019	-0.019	-0.016	-0.003
	(0.013)	(0.013)	(0.011)	(0.012)
Treatment Group	-0.005	0.005	0.014*	0.008
	(0.009)	(0.009)	(0.008)	(0.008)
Year 2005	-0.016**	0.016**	0.009	0.007
	(0.008)	(0.008)	(0.006)	(0.007)
Controls	Yes	Yes	Yes	Yes
Adj.R-Squared	-0.241	-0.241	-0.288	-0.228
Observations	2838	2838	2838	2838

Source: Own estimates based on school census data. Note: Robust standard errors in brackets, significance levels: * p<0.10, ** p<0.05, *** p<0.01

[21] World Bank (2001), however, notes that this distinction in contracted working hours is largely artificial.

Appendix 2.3 Considering Treatment Anticipation Effects

One of the concerns for the consistency of treatment effects estimated in this paper is anticipation of treatment. Considering the timeline of events outlined in Figure 2-2, schools file a budget proposal, possibly requesting a second teacher, in May. If the students can anticipate that they will receive a second teacher in the next year this may already influence end-of-year outcomes in December of the present school year; treatment anticipation could influence their perception of school quality and increase their willingness to complete. It is in this case inappropriate to include 2004 end-of-year outcomes in the matching procedure.

Table 2-12 shows the results for a matched difference-in-difference estimation for completion and failure shares, equivalent to Table 2-9; the only difference is the exclusion of the withdrawal and non-promotion share in 2004 among matching variables. In this scenario, it is possible to see a *treatment group effect*, i.e., a significant estimate of β_2 in equation (1).

Table 2-12. Difference-in-Difference Estimates for Alternative Matching Procedure

	[1]	[2]	[3]	[4]	[5]
Dependent Variable	Enrollment	Completed	Share Completed	Share Non-Promoted	Share Withdrawn
Treatment*2005	5.378***	3.704***	-0.009	-0.006	0.014
	(1.318)	(1.072)	(0.012)	(0.009)	(0.010)
Treatment Group	0.030	1.004	0.027***	-0.004	-0.023***
	(0.902)	(0.757)	(0.009)	(0.006)	(0.007)
Year 2005	-2.896***	-1.697**	0.012	-0.001	-0.010*
	(0.810)	(0.659)	(0.008)	(0.005)	(0.006)
Controls	Yes	Yes	Yes	Yes	Yes
Adj.R-Squared	0.200	0.149	0.099	0.111	0.075
Observations	2846	2846	2846	2846	2846

Note: p.p. stands for percentage points.

Column 1 shows that the enrollment effect is similar to the baseline results in the main section. There is a positive and strongly significant treatment effect on enrollment in the second year, stemming from reduced between-year dropout in treated schools.[22] This also results in a higher level of completion in the second year, induced by treatment (column 2). In contrast to the previous matching procedure, however, columns 2 and 3 indicate that there is indeed a treatment group effect. Given similar enrollment levels and lagged end-of-year outcomes, treated schools have a higher completion rate of about 2.7 percentage points already in 2004 (column 3); this is equivalent to about one more successful student. This effect stems mainly from fewer withdrawers *before* treatment, as indicated in column 5. Comparably to before, we still cannot reject the hypothesis that there is a significant after-treatment effect on the completion rate in the second year. Here, the effect is even slightly negative. Standard errors are not large in the estimates: an after-treatment effect of about 2.4 percentage points, i.e., about one more completed student, would be significant.

There are several possible explanations for this phenomenon. It is indeed possible that the treatment group effect represents an anticipation effect of receiving an additional teacher in treated schools, such that students may be less inclined to leave school. Possibly, they perceive that a quality increase will raise their utility from schooling, e.g., through higher labor market returns from education. Treatment anticipation would thus increase the share of grade completers even before treatment, especially via reduced withdrawal. One possible explanation for this is that the anticipation of treatment does not increase learning achievement itself and thus the probability to be promoted. Instead, it may increase expected future school quality and thus induce fewer children to leave presently. This would partly be in line with Hanushek at al. (2006) who find that high-quality schools keep students in school longer, even though in this case input quantity is raised, not output quality.

Nevertheless, a competing explanation is that this matching procedure does not properly account for the heterogeneity between treated and untreated schools which affects both treatment allocation and educational outcomes at the school level. At least two examples are realistic: First, treated and untreated schools could be on a different trajectory already before treatment. For instance, a new director may be responsible in

[22] All of the supplementary tables (not displayed) show the same patterns as for the results in the main section. The enrollment effect is spread out over all grades and stems mainly from keeping more students within the same school between years.

2004 both for applying for a new teacher as well as pushing more students to complete the grade. Second, the end-of-year outcomes of 2004 may influence the decision of whether or not an additional teacher is allocated to the single-teacher schools. In both cases, the observed treatment group effect would not result from treatment anticipation but be an artifact of inappropriate matching.

Even though it cannot be resolved which of the explanations is true, they do not change the main finding: a significant reduction in the student-teacher ration in rural single-teacher schools in Peru does not induce significantly higher completion rates. Raising teacher quantity is thus unlikely to solve Peru's problems of educational inefficiency.

Chapter 3

3 The Impact of Teacher Subject Knowledge on Student Achievement: Evidence from Within-Teacher Within-Student Variation

3.1 Introduction

One of the biggest puzzles in educational production today is that while there is clear evidence that teacher quality is a key determinant of student learning, little is known about which specific observable characteristics of teachers can account for this impact (e.g., Rockoff 2004; Rivkin et al. 2005).[23] In particular, there is little evidence to suggest that those characteristics most often used in hiring and salary decisions, namely teachers' education and experience, are crucial for teacher quality. Virtually the only characteristic that has been shown to be more frequently significantly correlated with student achievement is teachers' academic skills measured by scores on achievement tests (for reviews, cf. Wayne, Youngs 2003; Eide et al. 2004; Hanushek, Rivkin 2006). The problem with the latter evidence, however, is that issues of omitted variables and selection bias are intricately hard to address when it comes to estimating causal effects of teacher

[23] This chapter is based on joint work with Ludger Wößmann.

characteristics. Ask any parent about their children's teachers' quality, and it is immediately obvious that most are well aware of who is a good or bad teacher based on traits that generally remain unobserved by researchers. It is thus all but random which parents manage to get their children into the classrooms with the best teachers, both across and within schools.

In this paper, we provide estimates of the impact of teachers' academic skills on students' academic achievement that circumvent problems from unobserved teacher traits and non-random sorting of students to teachers using a unique data set from Peru. It contains test scores in two different academic subjects not only for each student, but also for each teacher. This allows us to identify the impact of teachers' academic performance in a specific subject on students' academic performance in the subject, while at the same time holding constant any student characteristics and any teacher characteristics that are constant across subjects. We can observe whether the same student taught by the same teacher in two different academic subjects performs better in one of the subjects if the teacher's knowledge is relatively better in this subject. Thus, our models can identify the effect based on within-teacher within-student variation by controlling for student fixed effects, teacher fixed effects, and subject fixed effects. We can additionally restrict our analysis to small schools with at most one teacher per grade, excluding any remaining possibility that parents may have chosen a specific teacher for their students in both subjects based on the teacher's specific knowledge in one subject.

Our identification strategy for effects of teacher characteristics is an extension of the within-student comparisons in two different subjects proposed by Dee (2005; 2007) which are capable of holding constant any student characteristics. His approach, however, is deemed to use variation across different teachers because the teacher characteristics that he analyses – gender, race, and ethnicity – do not vary within teachers. As a consequence, the identifying variation may still be related to issues of selection and unobserved teacher characteristics. By contrast, we are able to extend his identification strategy to use variation not only within individual students but also within individual teachers, because subject knowledge does vary within individual teachers and because our dataset allows us to observe teachers' knowledge in two different subjects. The teacher test results were evaluated using Rasch modeling and are thus convincing proxies of actual teacher knowledge in the subject.

We find that a one standard deviation increase in teacher test scores increases student test scores by about 4 percent of a standard deviation. Due to measurement error in test scores our result is attenuated downwards. A sensitivity analysis with plausible teacher test reliability measures shows that the effect is easily much larger, for example about 7 percent of a standard deviation when applying the student person separation reliability measures. Extended analysis using different sub-samples and quantile regressions does not indicate that there are strong patterns of non-linearity in the impact. We also find that the main result holds for the Peruvian 6th grade student population at large. The results should be understood in a developing country context of overall low schooling quality and very heterogeneous student and teacher skills. The analysis implies that in a developing country teacher subject knowledge is a key determinant of student learning and that teacher training should be high on the educational policy agenda.

3.2 Background and Literature Review

Our contribution is to provide a well-identified estimate of the impact of teacher subject knowledge on student achievement, employing a comprehensive measure of teacher subject knowledge while at the same time circumventing problems of selection and omitted variable bias. The following literature review details the context in which this study should be seen.

The empirical literature on the determinants of student learning has been tackling the issue of the impact of teachers from two sides: measuring the impact of teacher quality as a whole versus measuring the impact of distinct teacher characteristics. While the importance of teachers in the process of students' skill formation is undisputed it is econometrically challenging to isolate this effect.

Due to recent breakthroughs in the first stream of literature, researchers have firmly established that overall teacher quality is very heterogeneous, i.e., that teachers vary strongly in their impact on student outcomes (cf. Rivkin et al. 2005; Rockoff 2004). The key papers estimate total teacher quality effects focusing on changes in student achievement. It is only recently that there exist longitudinal data sources rich enough to allow convincing estimates. Rockoff (2004, pp. 247–248) concludes for the US school system that a "one-standard-deviation increase in teacher quality raises test scores by approximately 0.1 standard deviations in reading and math on nationally standardized distributions of achievement." A similar effect magnitude is found by Rivkin et al (2005).

The second stream of literature examines which specific teacher characteristics are responsible for these big effects and thus constitute the unobserved bundle of overall teacher quality associated with a certain person. Answers to this question are important for educational policy making since they could change the way we think about tying hiring policies and compensation schemes to specific teacher characteristics. The literature on educational production functions and attempts to resolve these questions dates back to the US government study "Equality of Educational Opportunity" (Coleman et al. 1966). Since then, several hundred studies have estimated traditional parametric educational production functions to calculate the effects of teacher characteristics and other educational inputs on student achievement.

In the empirical examination of different teacher attributes like teacher education, experience, salaries, test scores and certification, only teacher knowledge measured by teacher test scores has more consistently been associated with positive student outcomes. Already Coleman et al. (1966) found that verbal skills of teachers were associated with higher student learning. Synthesizing a growing amount of scientific literature on the issue, Hanushek (1986, p. 1164) concludes that "[t]he closest thing to a consistent finding among the studies is that 'smarter' teachers, ones who perform well on verbal ability tests, do better in the classroom, but even for that the evidence is not very strong." A decade later, Hanushek counts a total of 41 estimates of the effect of teacher test scores and finds that "[o]f all the explicit measures [of teachers and schools] that lend themselves to tabulation, stronger teacher test scores are most consistently related to higher student achievement, even though only 37% provide positive and statistically significant effects." (Hanushek 1997, p. 144). Eide et al. (2004, p. 233) suggest that "a stronger case can be made for measures of teachers' academic performance or skill as predictors of teachers' effectiveness."

Nevertheless, on all of these teacher characteristics, including teacher test scores, the evidence is far from conclusive, and many studies lack persuasiveness due to insufficient econometric methods (cf. Hanushek, Rivkin 2006). In the absence of high-quality data sets on matched student-teacher pairs over time, results have been largely plagued by econometric problems: omitted variables, sorting of students with different levels of knowledge between schools of different quality, and sorting within schools into classrooms with peers and teachers of differing quality. They can only be solved with comprehensive data and/or clever identification strategies. In their survey, Hanushek and Rivkin (2006) separate high and low-quality studies that estimate the association between teachers' scores on different tests of academic achievement and the gains in student achievement.[24] They note that "while the evidence is stronger than that for other explicit teacher characteristics, it is far from overwhelming" (p. 1064). Also, teacher academic skills can only explain a small portion of the overall variation in teacher impact on student achievement.

Our knowledge from teacher test score studies is hard to evaluate, not only because of the mentioned econometric problems but also because the form and scope of tested skills varies greatly. The types of employed scores are mostly verbal ability, or a mixture of different subjects from college entrance or teacher licensure examination scores. What is common among the studies is that the tested teacher skills can only be weakly tied to the academic knowledge in the subject in which student achievement is examined. The test score may be measured with considerable measurement error, e.g., when using the result from one single math question as an indicator of math skills (cf. Rowan, Chiang, Miller 1997). Furthermore, the examined skill may not be subject-specific, such as verbal ability (cf. Coleman et al. 1966). Also, the test scores may reflect an aggregate of different skills, such as verbal and pedagogic ability, English, math and science knowledge and other components used in the U.S. National Teacher Examination (cf. Summers, Wolfe 1977). While results from these tests are clearly interesting they cannot distinguish between subject-specific cognitive and general non-cognitive teacher skills.

[24] Important studies estimating the association of teacher test scores with student achievement gains include Hanushek (1971; 1992), Summers and Wolfe (1977), Murnane and Phillips (1981), Ehrenberg and Brewer (1995), Ferguson and Ladd (1996), Rowan, Chiang, and Miller (1997), and Ferguson (1998); see Hanushek (1997), Wayne and Youngs (2003), and Eide, Goldhaber, and Brewer (2004) for reviews.

To our knowledge, no study so far uses subject-specific teacher academic skill scores to explain student performance in the same subject while controlling for student and teacher fixed effects. The credibility of our estimates is enhanced as teacher subject knowledge distributions in two subjects are generated using scientific Rasch modeling.

3.3 Empirical Identification

3.3.1 Estimation Strategy

We specify an educational production function (1) with an explicit focus on teacher characteristics:

(1) $y_{is} = c_s + \alpha Z_i + \lambda X_{is} + \beta T_{is} + \gamma M_i + \varepsilon_{is}$

where test score y of student i in subject s is a function of

- a subject-specific constant c,
- non subject-specific student, classroom and school characteristics Z,
- subject-specific student, classroom and school characteristics X_s, including previous student subject knowledge,
- subject-specific teacher characteristics T_s, such as subject knowledge,
- non subject-specific teacher characteristics M, such as motivation, and
- a mean-zero error term ε_s.

The coefficient vectors β and γ characterize the impact of all subject-specific and non subject-specific teacher characteristics which constitute the overall teacher quality effect as estimated by Rivkin et al. (2005) and Rockoff (2004). Estimating this equation directly in order to characterize the impact magnitude of teacher subject knowledge would suffer from endogeneity problems due to unobserved factors, e.g., if previous student subject knowledge is not observed and teachers are non-randomly allocated to students according to previous subject knowledge.

A possible solution to get rid of fixed effects are within-student comparisons in two different subjects proposed by Dee (2005; 2007) which are capable of holding constant any student characteristics. The differenced equation reads

(2) $\Delta y_i = \Delta c + \lambda\ \Delta X_{is} + \beta\ \Delta T_{is} + \Delta \varepsilon_i$

where Δ denotes the difference between subject-specific variables in two subjects, in our case mathematics and reading. Such equations were introduced by Dee (2005, 2007) and similarly estimated for example by Ammermüller, Dolton (2006) and Clotfelter, Ladd and Vigdor (2007) to determine if specific teacher attributes such as gender, ethnicity or credentials and their interaction with student characteristics would have an effect on student performance and teacher perceptions.

The remaining problem of this approach is a possible bias if teacher assignment is non-random with respect to the students' subject-specific propensity for achievement. This is the case if students are allocated to different classrooms for every subject according to their subject-specific knowledge status, and teachers are assigned according to their subject-specific capabilities, with different teachers for different classrooms. A possible solution to this problem would be to examine only the sub-sample of students with the same teacher in both subjects. However, this will eliminate all teacher effects if there is no variation in subject-specific attributes *within* the same teacher.

Our data set is unique to the extent that it provides exactly that: within-teacher variation for the same teacher in subject-specific knowledge. Teachers in the EN 2004 took tests on the same subjects in which students were tested, so that there is a measure of cognitive skills which varies by subject. For the sub-sample of students that are taught by the same teacher in both subjects, we can isolate the effect of subject-specific teacher knowledge by differencing out all attributes that do not vary by subject, e.g., motivation. We call this reduced sample "same-teacher sample".

The Peruvian setting allows us to additionally restrict our analysis to small schools with at most one teacher per grade, which can further exclude any remaining possibility that parents may have chosen a specific teacher for their students in both subjects based on the teacher's specific knowledge in one subject. We call this twice reduced sample "same-teacher one-classroom (STOC) sample".

When estimating (2), we can get consistent estimates of β, the causal impact of teacher subject knowledge on student subject knowledge (independent of the subject) under the following assumptions:

1. Subject-specific student, teacher, classroom and school characteristics have the same effect in math (M) and reading (R), i.e., $\lambda_M = \lambda_R = \lambda$ and $\beta_M = \beta_R = \beta$

2. Non-specific student, teacher, classroom and school characteristics have the same effect in both subjects, i.e., $\alpha_M = \alpha_R = \alpha$ and $\gamma_M = \gamma_R = \gamma$.

3. Zero correlation of ΔT_{is} and the error term, i.e., $E(\Delta T_{is}'\Delta \varepsilon_i)=0$.

4. No measurement error in included variables.

Please refer to Appendix 3.1 for a discussion of the assumptions including estimation results for an alternative specification. Refer to Appendix 3.2 for a discussion of possibly non-linear effects of teacher knowledge on student knowledge, and according estimation results using quantile regressions.

3.3.2 Measurement Error Correction

We expect our teacher test score estimates to be smaller than the true effect because of attenuation bias since teacher test scores are measured with error.[25] If we assume classical measurement error in teacher test scores we can postulate $T_{is} = T_{is}^* + e_{is}$, where T_s is the observed test score in subject s, T_s^* is the true test score, and e_s is an additive white noise measurement error with $E(e_s)=0$ and $E(T_s^*,e_s)=0$ for all subjects s and individuals i. In a level regression of student test score in one subject on teacher test score in the same subject this will lead to the well-known attenuation bias of OLS, where

$$\text{plim}\,\hat{\beta} = \beta\left(\frac{Var(T_i^*)}{Var(T_i^*)+Var(e_i)}\right)$$, and

$$\lambda = \left(\frac{Var(T_i^*)}{Var(T_i^*)+Var(e_i)}\right)$$ is often called reliability ratio.

[25] As we only identify the effect of current, not accumulated, teacher subject knowledge on test scores this will further understate the effect we estimate.

As teacher test scores are correlated across subjects first differencing will reduce the signal-to-noise ratio and attenuate the estimated test score effect even more. The reliability ratio becomes

$$(3) \qquad \lambda = \left(\frac{Var(\Delta T_i^*)}{Var(\Delta T_i^*) + Var(\Delta e_i)} \right)$$

$$= \left(\frac{Var(T_M^*) + Var(T_R^*) - 2Cov(T_M^*, T_R^*)}{Var(T_M^*) + Var(T_R^*) - 2Cov(T_M^*, T_R^*) + Var(e_M) + Var(e_R) - 2Cov(e_M, e_R)} \right)$$

In order to present a sensitivity analysis of the magnitude of attenuation bias in our estimates we need to assess the size of (3). It is reasonable to assume that test-taking measurement error in two unrelated subjects is random such that there is no correlation of errors across subjects, $Cov(e_M, e_R) = 0$, or correlation between observed test score and measurement error across subjects, $Cov(T_M, e_R) = Cov(e_M, T_R) = 0$. Yet, if teachers tend to be either knowledgeable or not, we expect positive correlation of true teacher test scores. Given these assumptions,

$$Cov(T_M^*, T_R^*) = Cov(T_M - e_M, T_R - e_R)$$

$$= Cov(T_M, T_R) - Cov(T_M, e_R) - Cov(e_M, T_R) + Cov(e_M, e_R) = Cov(T_M, T_R)$$

We can thus approximate the size of attenuation bias if we estimate the variance of true test scores, the variance of measurement error, and the covariance of observed test scores. We can get an idea of these numbers with help of reliability measures of EN 2004.

3.3.3 Data

This paper uses data from the 2004 Peruvian national evaluation of student achievement, the "Evaluación nacional del rendimiento estudiantil" (EN 2004). The sample of 6th graders covers more than 12,000 students from more than 800 randomly sampled primary schools. The sample is representative at the national level and for comparisons of urban versus rural areas, public versus private, and complete versus multi-grade[26] schools.

We use test data from EN 2004 in two subjects, reading and mathematics, which are separately taught subjects in the students' curriculum. As a unique feature of EN 2004, not only students but also teachers were required to take tests in their respective subject. Both sets of cognitive skill tests were evaluated by the unit for quality measurement (UMC) of the Peruvian ministry of education (MINEDU) using Rasch modeling.

We focus our analysis on students who are served by the same teacher in math and reading, and particularly those in schools with only one classroom in 6th grade. This combination is quite frequent as the Peruvian school system is characterized by many small, remote schools to serve the dispersed population. Table 3-1 summarizes student test scores scaled to a mean of 0 and standard deviation of 1 for this sub-group.

Table 3-1. Summary Statistics Student Test Scores – Same-Teacher One-Classroom Sample

	Share	Score Math		Score Reading		N
Same Teacher & One Classroom	1.000	0.00	(1.00)	0.00	(1.00)	4302
Urban Area	0.398	0.55	(0.92)	0.59	(0.89)	2295
Rural Area	0.602	-0.36	(0.88)	-0.39	(0.87)	2007
Public School	0.917	-0.10	(0.94)	-0.10	(0.96)	3728
Private School	0.083	1.14	(0.88)	1.08	(0.82)	574
Multigrade School	0.551	-0.37	(0.87)	-0.39	(0.86)	2003
Complete School	0.449	0.46	(0.96)	0.48	(0.94)	2299
Student 1st Language: Spanish	0.835	0.16	(0.95)	0.15	(0.96)	3735
Student 1st Language: Native	0.161	-0.78	(0.83)	-0.76	(0.83)	537
Male Student	0.512	0.09	(1.00)	0.02	(0.97)	2232
Female Student	0.488	-0.10	(0.99)	-0.02	(1.03)	2070

Note: Summary statistics calculated using sampling weights from EN 2004. Means in left columns, standard deviations in brackets.

[26] Multi-grade schools are a distinct feature of many developing countries where parts of the population live in sparsely populated areas. For example, the remoteness of communities in the Andes and the Amazon basin makes it difficult to appropriately serve many students in rural Peru. Such places suffer from a small supply of qualified teachers and lack of critical student mass. As a result, multi-grade schools are a wide-spread phenomenon where several grades are served in the same class by the same teacher (cf. Hargreaves et al. 2001).

The "same-teacher one-classroom" (STOC) sub-sample still contains more than 4000 observations. Boys and girls score very similarly in reading while boys seem to fare better in math (0.09 compared to -0.10). All other subdivisions in groups produce apparently more different results between groups in both subjects: students achieve higher scores in rural compared to urban schools, in private compared to public schools and in complete compared to multi-grade schools. Spanish-speaking students fare significantly better than native Peruvians.

Table 3-2 summarizes teacher characteristics in the STOC sample. While male and female teachers score similarly in math tests, male teachers are much worse in reading than female teachers (-0.19 compared to 0.28). The quality of staff reflects the same pattern as the quality of students from Table 3-1. Teachers are better on average in both subjects in urban compared to rural areas, in private compared to public schools and in complete compared to multi-grade schools. Teachers with a university degree, compared to an institute, score worse in math (-0.36 compared to 0.10) and better in reading (0.10 compared to -0.02). Teachers in their first year with a class are worse in reading but not in math than those with more years with the same class.

Table 3-2. Summary Statistics Teacher Test Scores – Same-Teacher One-Classroom Sample

	Share	Score Math	Score Reading	N
Same Teacher & One Classroom	1.000	0.00 (1.00)	0.00 (1.00)	335
Male Teacher	0.577	0.01 (0.99)	-0.19 (0.92)	163
Female Teacher	0.423	0.03 (0.99)	0.28 (1.04)	149
Urban Area	0.408	0.20 (0.88)	0.23 (0.91)	121
Rural Area	0.592	-0.14 (1.06)	-0.15 (1.03)	214
Public School	0.914	-0.04 (0.99)	-0.03 (0.98)	299
Private School	0.086	0.42 (1.02)	0.35 (1.15)	36
Multigrade School	0.544	-0.14 (1.03)	-0.19 (0.94)	228
Complete School	0.456	0.17 (0.94)	0.23 (1.03)	107
Teacher Degree: University	0.239	-0.36 (0.99)	0.10 (1.24)	91
Teacher Degree: Institute	0.761	0.10 (0.98)	-0.02 (0.92)	237
Teacher 1 Year with Class	0.408	0.04 (1.00)	-0.14 (0.89)	127
Teacher 2 Years with Class	0.318	0.03 (1.00)	0.18 (1.11)	90
Teacher 3-6 Years with Class	0.274	0.02 (1.03)	0.11 (0.96)	70

Note: Summary statistics calculated using sampling weights from EN 2004. Means in left columns, standard deviations in brackets. The number of observations of collectively exhaustive sub-samples does not always add up to the full sample due to missing information in the sub-group characteristics.

Table 3-3 summarizes subject-specific control variables in estimating equation (1). They account for motivational differences of the student between subjects, differences in teaching hours and differences in curriculum design by the teacher.[27] These differences may be problematic if they are systematically correlated with the difference in teacher subject knowledge between subjects. For example, students may be more motivated in subjects with more knowledgeable teachers, or teachers may prefer to teach subjects in which they are more knowledgeable. In all respects, there are no systematic differences between subjects.

Table 3-3. Summary Statistics Control Variables – Same-Teacher One-Classroom Sample

	Math		Reading	
	Mean (S.D.)	N	Mean (S.D.)	N
Motivation Index	4.27 (0.84)	4302	3.72 (1.06)	4302
Teaching Hours	5.91 (1.60)	3855	5.92 (1.63)	3822
Curriculum Design: Subject-specific Books	0.69 (0.47)	4140	0.62 (0.49)	4213
Curriculum Design: Student Working Books	0.52 (0.50)	4140	0.49 (0.50)	4213
Curriculum Design: Local School Guidelines	0.56 (0.50)	4140	0.57 (0.50)	4213
Curriculum Design: Institutional Guidelines	0.24 (0.43)	4140	0.31 (0.46)	4213
Curriculum Design: Regional Guidelines	0.02 (0.14)	4140	0.01 (0.12)	4213
Curriculum Design: National Guidelines	0.80 (0.40)	4140	0.80 (0.40)	4213
Curriculum Design: Adj. Curriculum Guidelines	0.34 (0.48)	4140	0.33 (0.47)	4213
Curriculum Design: Others	0.16 (0.37)	4140	0.14 (0.35)	4213
Curriculum Design: None	0.01 (0.10)	4140	0.01 (0.10)	4213

Note: Summary statistics calculated using sampling weights from EN 2004. Means in left columns, standard deviations in brackets.

The same-teacher one-classroom sample is not a representative sample of the Peruvian student population: predominantly multi-grade schools, rural schools and public schools employ only one teacher in 6th grade to teach both math and reading to the same classroom. Appendix 3.3 summarizes the characteristics of the full student sample. Also,

[27] The student motivation index from 0-5 for both subjects is calculated from 5 survey questions corresponding to each subject such as "I like to read in my free time" (reading) or "I have fun solving mathematical problems" (math) in which students can decide between disagree/sometimes true/agree. The resulting answer is coded as 0 for the answer which displays low motivation for the subject, 0.5 for medium motivation, and 1 for high motivation.

it contains extended results which support the argument that estimation results obtained for the STOC sub-sample are representative for the Peruvian student population at large.

3.4 Results

3.4.1 Main Results

Table 3-4 presents the results from estimating specifications (1) and (2) without control variables.[28] All estimates are calculated using robust standard errors clustered at the classroom level. The first column contains results from naïve OLS estimation in the full EN 2004 sample with pooled subjects which contains around 40 percent students with different teachers in math and reading. The estimated coefficient of teacher test score is significantly positive at 0.24 and reflects the strong bias from between-school student sorting. Column 2 estimates the same specification including student fixed effects. The teacher test score coefficient drops to an insignificant small effect and also reflects biases. For example, weaker students may be allocated to better teachers which could mitigate or completely cancel a positive teacher test score effect.

Table 3-4. Regression Results – Introducing Fixed Effects

Estimating Method	OLS	OLS	1st Diff.	1st Diff.
Sample	Full	Full	Same Teacher	STOC
Teacher Score	0.235***	0.012	0.037***	0.045***
	(0.021)	(0.009)	(0.012)	(0.014)
Student Fixed Effects	No	Yes	Yes	Yes
Teacher Fixed Effects	No	No	Yes	Yes
Adj.R-Squared	0.058	0.000	0.002	0.004
F-Statistic	121.0	1.9	8.9	10.9
Observations (Students)	12165	12165	6819	4302

Note: Robust standard errors in brackets, observations clustered at classroom level; * p<0.10, ** p<0.05, *** p<0.01.

[28] All regression models are fitted without sampling weights since we want to estimate a structural relationship and because our weights were determined due to stratification with respect to independent variables (cf. Winship, Radbill 1994). We thus fit all models with a constant because the unweighted differenced student test score is not zero.

Columns 3 and 4 of Table 3-4 show the results of estimating the first differenced equation (2) which implicitly includes student fixed effects. By restricting the student sample to those with the same teacher, we reduce possible bias stemming from within-school sorting and implicitly take care of teacher fixed effects. In both cases there is a positive and highly significant effect of teacher subject knowledge on student achievement. The third column displays estimates for the same-teacher sample: an increase in teacher test scores of one standard deviation increases student test scores by about 3.7 percent of a standard deviation. The fourth column shows the result for estimation in the same-teacher one-classroom (STOC) sample which excludes any remaining possible bias from sorting of students within the grade. The coefficient magnitude is very similar at 4.5 percent of a standard deviation and statistically not distinguishable from the third column.

Table 3-5 presents the key results from specification (2) using the same-teacher and the same-teacher one-classroom sample and subsequently adding several control variables. The first columns repeat the main findings from column 3 and 4 of Table 3-4.

Table 3-5. First Differencing Results with Control Variables

Panel A. Same Teacher Sample

Dep. Var. Student Score Difference M-R	[1]	[2]	[3]	[4]
Teacher Score Difference M-R	0.037***	0.037***	0.031**	0.027*
	(0.012)	(0.012)	(0.014)	(0.014)
Student Motivation Difference		0.021**	0.020**	0.019**
		(0.009)	(0.009)	(0.009)
Teaching Hours Difference			0.031***	0.029***
			(0.011)	(0.010)
Constant	-0.040***	-0.053***	-0.048***	-0.051***
	(0.013)	(0.014)	(0.015)	(0.016)
Controls Teaching Method Difference	No	No	No	Yes
Adj.R-Squared	0.002	0.003	0.004	0.003
F-Statistic	8.9	7.1	5.9	1.9
Observations	6819	6819	6010	5769

Panel B. STOC Sample

Dep. Var. Student Score Difference M-R	[1]	[2]	[3]	[4]
Teacher Score Difference M-R	0.045***	0.045***	0.037**	0.031**
	(0.014)	(0.014)	(0.015)	(0.015)
Student Motivation Difference		0.013	0.012	0.011
		(0.011)	(0.012)	(0.013)
Teaching Hours Difference			0.026**	0.033***
			(0.012)	(0.013)
Constant	-0.029*	-0.036**	-0.031	-0.031
	(0.016)	(0.017)	(0.019)	(0.019)
Controls Teaching Method Difference	No	No	No	Yes
Adj.R-Squared	0.004	0.004	0.003	0.002
F-Statistic	10.9	6.0	4.2	1.4
Observations	4302	4302	3745	3592

Note: Robust standard errors in brackets, observations clustered at teacher level; * p<0.10, ** p<0.05, *** p<0.01

The second column introduces student motivation as an additional regressor which may partly be a result of the quality of teaching in the respective subject. Controlling for this effect does not reduce the magnitude of the teacher test score impact in either sample.

The third column introduces weekly teaching hours in mathematics and reading as an additional regressor to control for the possibility that teachers may prefer to teach more in the subject they know better, or get to know the subject better which they teach more. The effect of teaching hours is positive and significant. Increasing teaching time by one weekly hour increases student test scores by about 3 percent of a standard deviation. Apparently, the coefficient magnitude of teacher test scores is reduced. However, this effect comes fully from reducing the sample by several hundred observations which do not have teaching hour information.[29] This evidence is not suggestive of the fact that teachers who are better in one subject may spend more time and effort in that subject relative to the other one.

The same is true for the fourth column. Introducing controls for differences in the subject curriculum[30] apparently reduces the teacher test score effect. Again the reduction

[29] In Table 3-5, panel A, the estimated teacher test score effect using the specification of column [1] on the sample of column [3] is 0.031**; in panel B it is 0.036**.

[30] EN 2004 asks teachers to describe which items they use to design the subject curriculum: working books, school curriculum, institutional educational projects, regional educational projects, 3rd cycle basic curriculum

comes from reducing the sample by several hundred observation compared to column 1.[31] As we do not know the process which caused these missing observations we confide in our original estimates. We thus consider the effect magnitude of 0.037 and 0.045 robust to adding relevant control variables which may be correlated with teacher test scores.

3.4.2 Measurement Error Correction

In order to approximate the size of attenuation bias we need the covariance of observed test scores, the variance of unobserved true test scores and the variance of measurement error. The first can be observed (0.44 in the STOC sample), the second or third needs to be assessed to calculate the remaining figure. EN 2004 data provides estimates of psychometric properties, however, only of *student* test scores. The provided measure is *person separation reliability*, the proportion of observed sample variance that is not attributable to measurement error (Wright, Masters 1982) which measures how reliably test-takers of high and low ability can be distinguished. The definition corresponds to the reliability measure

$$\lambda = \left(\frac{Var(T_i^*)}{Var(T_i^*) + Var(e_i)} \right)$$

that attenuates the regression results. Knowing teacher test reliability would thus allow calculating the size of attenuation bias. 6th grade student test reliability is 0.72 for reading and 0.85 for math. Person separation reliability increases with the length of the test and the ability range within the sample. Since teachers took a shorter test than students but their ability range is possibly larger we cannot predict their test reliability compared to student test reliability and thus present a sensitivity analysis of attenuation bias size depending on reliability measures.

---

structure and/or readjusted curricular programs. When using each of them as differenced dummy in the regression only one becomes weakly significant and they are jointly insignificant.

[31] In table 5a the estimated teacher test score effect using the specification of column [1] on the sample of column [4] is 0.028*, in table 5b it is 0.032**.

Table 3-6. Measurement Error Correction for STOC-sample

	Assumption about Person Separation Reliabilities in Teacher Tests						
	No Error	= Student Test	Range of Plausible Reliabilities				
Assumed Reliability - Math	1.00	0.85	0.90	0.80	0.70	0.60	0.50
Assumed Reliability - Reading	1.00	0.72	0.90	0.80	0.70	0.60	0.50
Resulting Attenuation Factor	1.000	0.616	0.821	0.643	0.464	0.286	0.107
Corrected Teacher Test Score Effect	0.045	0.073	0.055	0.070	0.097	0.158	0.420

Note: Results calculated for the STOC sample, i.e., a covariance of observed teacher test scores of 0.44 and first differencing estimation result of 0.045. Since test scores are scaled to standard deviation and variance of 1, reliabilities correspond to true test score variances, and, given white noise error assumptions, error variances to 1 minus reliability.

Table 3-6 presents corrected Teacher Test Score Effects for different values of teacher test reliability when plugged into (3). The first column presents the results in absence of measurement error for the STOC sample, i.e., as if teacher subject knowledge was measured accurately by the test. The second column presents estimates as if the teacher test had the same reliability as the student test. For student reliability measures of 0.72 and 0.85 the teacher test score effect is attenuated by a factor of 0.62; the real effect would thus be 7.3 percent student test score increase of a standard deviation for a one standard deviation increase in teacher test score. The right panel presents a range of other plausible reliability estimates. For example, for a reliability of 0.7 in both tests, the teacher test score effect would be almost 10 percent.

3.4.3 Effects in Different Sub-Samples

Dividing the full sample into collectively exhaustive sub-samples shows the effect of teacher test scores on student achievement in different settings (see Table 3-7).

Table 3-7. Regression Results for Sub-Samples

Sample	Same-teacher			Same-teacher, one-classroom (STOC)		
	Beta	Rob. S.E.	N	Beta	Rob. S.E.	N
All	0.0371	(0.0125)	6819	0.0446	(0.0135)	4302
Urban Area	0.0314	(0.0162)	4666	0.0397	(0.0194)	2295
Rural Area	0.0482	(0.0182)	2153	0.0504	(0.0184)	2007
Public School	0.0373	(0.0120)	6054	0.0500	(0.0135)	3728
Private School	0.0377	(0.0580)	765	0.0024	(0.0498)	574
Multigrade School	0.0426	(0.0219)	2149	0.0523	(0.0217)	2003
Complete School	0.0334	(0.0153)	4670	0.0384	(0.0177)	2299
Student 1st Language: Spanish	0.0331	(0.0133)	6121	0.0400	(0.0147)	3735
Student 1st Language: Native	0.0714	(0.0356)	652	0.0777	(0.0370)	537
Male Student	0.0274	(0.0150)	3471	0.0395	(0.0169)	2232
Female Student	0.0415	(0.0169)	3348	0.0485	(0.0186)	2070
Male Teacher	0.0239	(0.0196)	2791	0.0185	(0.0217)	1942
Female Teacher	0.0339	(0.0176)	3623	0.0461	(0.0188)	2115
Student-Teacher Same Gender	0.0595	(0.0172)	3294	0.0724	(0.0187)	2022
Student-Teacher Diff. Gender	0.0138	(0.0162)	3120	0.0198	(0.0182)	2035
Teacher 1 Year with Class	0.0482	(0.0217)	2069	0.0655	(0.0236)	1478
Teacher 2 Years with Class	0.0541	(0.0219)	2355	0.0357	(0.0234)	1442
Teacher 3-6 Years with Class	0.0532	(0.0269)	1651	0.0953	(0.0303)	886
Teacher Degree: University	0.0399	(0.0221)	1810	0.0250	(0.0210)	1182
Teacher Degree: Institute	0.0370	(0.0155)	4787	0.0606	(0.0175)	2982

Note: Robust standard errors in brackets, observations clustered at teacher level.. Regression results for specification without control variables. Measurement error correction with reliability factor of 0.62

Even though the sample size reduces considerably in the different sub-groups, a positive, statistically and economically significant effect appears in almost all sub-samples. In the same teacher sample (left panel) it mostly varies between about 0.03 and 0.04, in the STOC sample (right panel) between about 0.04 and 0.05 when measurement error is not corrected for. The effect is small and statistically weak or insignificant in private schools, for student-teacher different gender pairs, male teachers and teachers with university degree.

We cannot reject the hypothesis that there is no effect of teacher test scores in private schools. Possible reasons are the small sample size of private school students, different learning transmission mechanisms in private than public schools or non-linearities in the teacher test score effect as private student tend to be in the high range of test scores in both subjects.

Comparing the results for students whose mother tongue is Spanish and native the point estimates suggest that there may be a bigger effect for native (0.071) than Spanish (0.033) students.

The most drastic difference between the effect estimate for two categories is between student teacher pairs of the same gender and those with different genders. While for the same-gender case the effect is estimated to be 0.073 and highly significant, the effect is 0.02 and insignificant for the different-gender case in the STOC sample. Only in this case can the effect in two mutually exclusive subsets be statistically distinguished.[32] This finding may suggest that in order for the teacher to transmit knowledge to students there must be some connection between the two which may be facilitated by sharing the same gender.

Extended results are presented in the appendices 1, 2 and 3. We estimate and discuss results which allow for subject-specific impacts of teacher knowledge (Appendix 3.1). We conclude that teacher test score effects cannot be consistently estimated without the assumption that effects do not vary between subjects. We also consider a possible non-linearity of effects by estimating quantile regressions (Appendix 3.2). Overall, the evidence from OLS regressions in different sub-samples and quantile regressions is not suggestive of the fact that there are strong non-linearities in the impact of teacher subject knowledge on student subject knowledge. Furthermore, we discuss the generalizability of results to the Peruvian student population at large (Appendix 3.3). Even though the same-teacher sample on which the main estimations are performed is not a representative sub-sample of the Peruvian student population, the results from a matching procedure indicate that the estimated teacher test score effects hold for the Peruvian student population at large.

3.5 Conclusion

We believe that this paper has presented a well-identified estimate of teacher academic skills on student achievement by exploiting within-teacher within-student

[32] Equality of coefficients was tested by regressing the difference in student test scores on dummy variables for the sub-sample (e.g., urban and rural) and interaction effects between sub-sample dummy variables and teacher test score difference. Afterwards, a t-test was conducted to test the equality of interaction effects.

variation in test scores. We find that a one standard deviation increase in teacher subject knowledge increases student achievement by around 4 percent of a standard deviation, and by more when correcting for measurement error attenuation. This effect is robust to most sub-samples of the data and representative for the Peruvian 6th grade population at large. It should be interpreted as an effect in a developing country setting of low academic standards overall.

Unfortunately, the data does not permit to calculate a credible output-based total teacher quality distribution. Thus we cannot answer the question what share of the total teacher quality effect is due to teachers' subject knowledge.

References

Ammermüller, Andreas, Dolton, Peter (2006). Pupil-Teacher Gender Interaction Effects on Scholastic Outcomes in England and the USA. ZEW Discussion Paper No. 06-060. Mannheim: Centre for European Economic Research.

Clotfelter, Charles T., Ladd, Helen F., Vigdor, Jacob L. (2007). Teacher Credentials and Student Achievement: Longitudinal Analysis with Student Fixed Effects. *Economics of Education Review, 26*(6), pp. 673–682.

Coleman, James S., Campbell, Ernest Q., Hobson, Carol F., McPartland, James M., Mood, Alexander M., Weinfeld, Frederic D., et al. (1966). *Equality of Educational Opportunity.* Washington, D.C.: Government Printing Office.

Dee, Thomas S. (2005). A Teacher Like Me: Does Race, Ethnicity, or Gender Matter? *American Economic Review, 95*(2), pp. 158–165.

Dee, Thomas S. (2007). Teachers and the Gender Gaps in Student Achievement. *Journal of Human Resources, 42*(3), pp. 528–554.

Ehrenberg, Ronald G., Brewer, Dominic J. (1995). Did Teachers' Verbal Ability and Race Matter in the 1960s? Coleman Revisited. *Economics of Education Review, 14*(1), pp. 1–21.

Eide, Eric, Goldhaber, Dan, Brewer, Dominic (2004). The Teacher Labour Market and Teacher Quality. *Oxford Review of Economic Policy, 20*(2), pp. 230–244.

Ferguson, Ronald F. (1998). Can Schools Narrow the Black-White Test Score Gap? In Christopher Jencks & Meredith Phillips (Eds.), *The Black-White Test Score Gap,* pp. 318–374. Washington, D.C.: Brookings Institution.

Ferguson, Ronald F., Ladd, Helen F. (1996). How and Why Money Matters: An Analysis of Alabama Schools. In Helen F. Ladd (Ed.), *Holding Schools Accountable: Performance-based Reform in Education,* pp. 265–298. Washington, D.C.: Brookings Institution.

Hanushek, Eric (1971). Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data. *American Economic Review, 61*(2), pp. 280–288.

Hanushek, Eric A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature, 24*(3), pp. 1141–1177.

Hanushek, Eric A. (1992). The Trade-Off between Child Quantity and Quality. *Journal of Political Economy, 100*(1), pp. 84–117.

Hanushek, Eric A. (1997). Assessing the Effects of School Resources on Student Performance: An Update. *Educational Evaluation and Policy Analysis, 19*(2), pp. 141–164.

Hanushek, Eric A., Rivkin, Steven G. (2006). Chapter 18 Teacher Quality. In Eric A. Hanushek & Finis Welch (Eds.), *Handbook of the Economics of Education.* Volume 2, pp. 1051–1078. Elsevier.

Hargreaves, Eleanore et al. (2001). Multigrade Teaching in Peru, Sri Lanka and Vietnam: an Overview. *International Journal of Educational Development, 21*(6), pp. 499–520.

Koenker, Roger W., Bassett, Gilbert, JR. (1978). Regression Quantiles. *Econometrica, 46*(1), pp. 33–50.

Murnane, Richard J., Phillips, Barbara R. (1981). What Do Effective Teachers of Inner-City Children Have in Common? *Social Science Research, 10*(1), pp. 83–100.

Rice, Jennifer K. (2003). *Teacher Quality: Understanding the Effectiveness of Teacher Attributes.* Washington, D.C.: Economic Policy Institute.

Rivkin, Steven G., Hanushek, Eric A., Kain, John F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica, 73*(2), pp. 417–458.

Rockoff, Jonah E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review, 94*(2), pp. 247–252.

Rowan, Brian, Chiang, Fang-Shen, Miller, Robert J. (1997). Using Research on Employees' Performance to Study the Effects of Teachers on Students' Achievement. *Sociology of Education, 70*(4), pp. 256–285.

Summers, Anita A., Wolfe, Barbara L. (1977). Do Schools Make a Difference? *American Economic Review, 67*(4), pp. 639–652.

Wayne, Andrew J., Youngs, Peter (2003). Teacher Characteristics and Student Achievement Gains: A Review. *Review of Educational Research, 73*(1), pp. 89–122.

Winship, Christopher, Radbill, Larry (1994). Sampling Weights and Regression Analysis. *Sociological Methods and Research, 23*(2), pp. 230–257.

Wright, Benjamin D., Masters, Geofferey N. (1982). *Rating Scale Analysis.* Chicago: Mesa Press.

Appendix 3

Appendix 3.1 Subject-specific Teacher Effects

In section 3.3.1, we specify an educational production function

(1) $y_{is} = c_s + \alpha\, Z_i + \lambda\, X_{is} + \beta\, T_{is} + \gamma\, M_i + \varepsilon_{is}$

and it first-differenced version

(2) $\Delta y_i = \Delta c + \lambda\, \Delta X_{is} + \beta\, \Delta T_{is} + \Delta\varepsilon_i$.

We name three functional assumptions in order to get consistent estimates of β, the causal impact of teacher subject knowledge on student subject knowledge (independent of the subject):

1. Subject-specific student, teacher, classroom and school characteristics have the same effect in math (M) and reading (R), i.e., $\lambda_M = \lambda_R = \lambda$ and $\beta_M = \beta_R = \beta$.

2. Non-specific student, teacher, classroom and school characteristics have the same effect in both subjects, i.e., $\alpha_M = \alpha_R = \alpha$ and $\gamma_M = \gamma_R = \gamma$.

3. Zero correlation of ΔT_{is} and the error term, i.e., $E(\Delta T_{is}'\Delta\varepsilon_i)=0$.

An example in which assumptions 1 and 2 do not hold would be if reading skills are mainly acquired at home, and mathematics skills mainly in school. As a consequence, school inputs such as teacher subject knowledge or non-specific teacher fixed effects (e.g. a teacher's overall motivation to teach or pedagogic ability) may have different impacts in different subjects. The correct specification of the differenced educational production function may thus be

(3) $\Delta y_i = \Delta c + \beta_M T_{iM} + \beta_R\, (-\, T_{iR}) + (\gamma_M - \gamma_R)\, M_i + \Delta\varepsilon_i$ where

$(\gamma_M - \gamma_R)\, M_i + \Delta\varepsilon_i = \nu_i$

Since M_i is unobserved this relates to the question if the third assumption holds. Even if there are left out subject-specific effects they will only harm the consistency of the estimated teacher knowledge effects under certain conditions. (i) To calculate β in (2) and β_M and β_R in (3) consistently there must be no systematic correlation between the difference in subject-specific unobservables and in subject-specific teacher test scores, i.e., $E(\Delta T_{is}'\Delta\varepsilon_i) = 0$. A concern for this could be the unobserved difference of prior achievement of students in the respective subjects. (ii) Additionally, in the case of (3), it must hold that $E(T_{iM}'(\gamma_M - \gamma_R)M_i)=0$ and $E(-T_{iR}'(\gamma_M - \gamma_R)M_i)=0$. A concern for this could be subject-specific effects of non-specific teacher skills.

i. A non-zero correlation of ΔT_{is} and Δe_i is possible if within one grade students are allocated on grounds of within-student performance differences between the subjects to appropriate teachers, i.e., students and teachers have correlated within-person knowledge differences. The possibility of this bias is eliminated when reducing the sample to those schools with only one class in grade 6 since in these schools student sorting is impossible. Thus, even if our data does not contain previous student subject test scores which are essential for cumulative educational production functions in non-differenced specifications, we are confident to eliminate any bias arising from lack of this information.

ii. If the impact of non-cognitive teacher skills varies by subject the regression error contains the term $(\gamma_M - \gamma_R)M_i$. Without loss of generality assume that $\gamma_M > \gamma_R$. It is likely that M_i is positively correlated with T_{iM} and negatively with $(-T_{iR})$, i.e., that non subject-specific teacher skills, such as motivation, are correlated with subject-specific teacher skills. In this case, β_M is biased upward and β_R downward. Thus, even if we observe different coefficients when estimating in (3), this may come from regression bias and not from subject-specific effect sizes.

    If we believe that $\beta_M = \beta_R = \beta$ but that $\gamma_M \neq \gamma_R$, the critical assumption for consistency of β is $E(\Delta T_{is}'M_i)=0$. There is no obvious reason to believe that the difference in subject-specific teacher skills is systematically related to the level of non subject-specific teacher skills.[33] The consistency of (2) is in this case not harmed even if $\gamma_M \neq \gamma_R$. As a consequence, we can consistently estimate (2) if assumption 2 fails but 1 holds, however, we cannot consistently estimate (3) if both assumptions 1 and 2 fail.

Results Allowing for Subject-specific Effects

Table 3-8 shows the results from estimating (3) in the same-teacher and STOC sample allowing for subject-specific teacher test score effects.

Table 3-8. Regression Results – Subject-specific Impacts

Dep. Var. Student Score Difference M-R	Same Teacher	STOC
Teacher Test Score (M)	0.065***	0.082***
	(0.017)	(0.021)
(–1) * Teacher Test Score (R)	0.025	0.032
	(0.015)	(0.022)
Constant	-0.032**	-0.035**
	(0.016)	(0.017)
Adj.R-Squared	0.005	0.007
F-Statistic	7.66	8.40
Observations	4302	3940

Note: Robust standard errors in brackets, observations clustered at teacher level;
* p<0.10, ** p<0.05, *** p<0.01

The estimated coefficients indicate that teacher mathematics knowledge may drive student test scores in mathematics much more than teacher reading knowledge drives student reading scores. The effects are statistically distinguishable. This may have two reasons: On the one hand, school inputs influence mathematics and reading differently, e.g., if the foundation for reading comprehension is laid out at home while mathematics is mainly taught in school. On the other hand, as argued in this appendix, teacher mathematics and reading scores separately may be correlated with other unobserved effects thus biasing estimated coefficients in the observed direction (mathematics biased

[33] The same line of argument should hold for other factors with possibly different impacts, e.g. household characteristics. It is unclear why they should be correlated with the difference in teacher subject-specific teacher knowledge (ΔT_{is}), but it is likely that they are correlated with overall teacher quality.

upwards, reading biased downwards). It is not possible to distinguish these two effects here.

Appendix 3.2 Non-linear Teacher Effects

Teacher academic skills may not uniformly matter to all students depending on their characteristics. For example, teacher knowledge may be especially important at very low levels of achievement, or contrarily, a minimum level of understanding may be necessary to gain from teacher knowledge. Hanushek and Rivkin (2006) note that "existing research gives no hints of whether there is any nonlinear impact of knowledge in different ranges." Rice (2003) suggests that teacher test scores may be particularly important for at-risk students.

When estimating (2), we impose a linearity assumption onto the relation between teacher knowledge and student achievement and estimate the rate of change in the mean of the student score distribution conditional on the teacher score. In order to allow for different effects in the different parts of the student test score distribution, it would be interesting to assess the structural relation by estimating quantile regressions, as proposed by Koenker and Bassett (1978). However, as quantiles are non-linear operators, contrarily to the expected value of OLS estimations, the interpretation of quantiles of the first difference is not directly transferable to quantile regressions in levels. We can estimate

(4) $\Delta y_i = \Delta c_\theta + \beta_\theta \Delta T_i + \Delta e_{\theta i}$

where β_θ is the vector of regression parameters associated with the θth percentile. However, the interpretation of the resulting parameters must be made in differences, i.e., results will only yield an explanation of the effect of differences in teacher subject knowledge on the conditional quantiles of the distribution of differences in student test scores between math and reading.

Quantile Regression Results

Table 3-9 shows the results for quantile regressions with bootstrapped standard errors of the differenced equation. As we mentioned above, the estimates thus refer to effects of teacher test score *difference* in different conditional quantiles on student test score *differences*. We can reject the hypothesis that there is no effect between the 20th and 90th percentile of the student score distribution. But even though the effect magnitude seems to vary in the different quantiles of the conditional distribution, standard errors do not allow us to reject the hypothesis that there is no difference in the effects in the different parts of the distribution. This means that there is no big difference between the impact of a teacher who scored relatively much higher in reading (lower quantiles) on his students' relative reading skills than the impact of a relatively better math teacher (upper quantiles) on his students' math skills.

Table 3-9. Quantile Regression Results

	Same Teacher			Same-teacher, one-classroom		
Percentile	Beta	S.E.	N	Beta	S.E.	N
10	0.0281	(0.0207)	6819	0.0327	(0.0231)	4302
20	0.0323	(0.0122)	6819	0.0373	(0.0130)	4302
30	0.0421	(0.0097)	6819	0.0449	(0.0115)	4302
40	0.0314	(0.0137)	6819	0.0467	(0.0124)	4302
50	0.0334	(0.0093)	6819	0.0442	(0.0117)	4302
60	0.0338	(0.0102)	6819	0.0471	(0.0117)	4302
70	0.0444	(0.0113)	6819	0.0541	(0.0104)	4302
80	0.0376	(0.0152)	6819	0.0590	(0.0170)	4302
90	0.0302	(0.0171)	6819	0.0431	(0.0229)	4302

Note: Standard errors in brackets, obtained by bootstrapping with 100 repetitions.

Appendix 3.3 Generalizability of Results

Table 3-10 shows how the same-teacher one-classroom sample is drawn from the whole school population: in comparison to the national average predominantly multi-grade schools, rural schools and public schools employ the same teacher to teach both math and reading to the same classroom and are small enough to have one classroom in 6^{th} grade only.

Table 3-10. Share of Classrooms in and out of STOC Sample by Type of School

Shares	Diiferent Teacher / >1 Classroom	Same-Teacher One-Classroom	All Teachers
Multigrade School	0.032	0.547	0.234
Rural Area	0.055	0.596	0.268
Public School	0.823	0.914	0.859
N	503	335	838

Note: Summary statistics calculated using sampling weights from EN 2004.

Since we pick the sub-sample of students with the same teacher in math and reading as our main sample for the analysis, you should keep in mind that this sample is predominantly drawn from schools which are disadvantaged in terms of the resources they can spend, and the pool of students and teachers from which they draw. Table 3-11 shows summary statistics for student test scores in math and reading which are scaled to mean 0 and standard deviation 1 for the full sample. Students from same-teacher one-classroom schools have significantly lower test scores in math and reading (-0.36 and -0.40) than their peers from other schools (0.34 and 0.36). Also, students from urban, private, and complete schools have higher test scores than their respective peers.

Table 3-11. Summary Statistics Student Test Scores – Full Sample

Students	Share	Score Maths	Score Reading	N
All	1.000	0.00 (1.00)	0.00 (1.00)	12165
Same Teacher in M and R	0.622	-0.23 (0.96)	-0.23 (0.98)	6819
Same Teacher & One Classroom	0.399	-0.36 (0.99)	-0.40 (1.01)	4302
Different Teacher in M and R	0.378	0.34 (0.93)	0.36 (0.90)	4990
Male Student	0.509	0.06 (1.01)	-0.02 (0.97)	6164
Female Student	0.491	-0.07 (0.99)	0.02 (1.03)	6001
Urban Area	0.729	0.23 (0.93)	0.26 (0.90)	9555
Rural Area	0.271	-0.63 (0.90)	-0.69 (0.92)	2610
Public School	0.861	-0.16 (0.92)	-0.15 (0.94)	9450
Private School	0.139	0.91 (0.95)	0.83 (0.87)	2359
Multigrade School	0.235	-0.63 (0.93)	-0.70 (0.95)	2446
Complete School	0.765	0.19 (0.94)	0.21 (0.91)	9719

Note: Summary statistics calculated using sampling weights from EN 2004. Standardization of test scores on full sample. The observation number of collectively exhaustive sub-samples does not always add up to the full sample due to missing information in the sub-group characteristics.

Table 3-12 summarizes teacher test scores in math and reading also scaled to mean 0 and standard deviation 1 for the full sample. The same patterns as for students hold: teachers lecturing both math and reading to the same students in one-classroom schools have considerably lower test scores than their specialized peers. Also, teachers of urban, private, and complete schools have higher subject knowledge than their respective peers.

Table 3-12. Summary Statistics Teacher Test Scores – Full Sample

	Maths			Reading		
	Share	Score	N	Share	Score	N
All	1.000	0.00 (1.00)	862	1.000	0.00 (1.00)	862
Same Teacher in M and R	0.625	-0.16 (0.94)	505	0.626	-0.13 (0.96)	505
Same Teacher & One Classroom	0.396	-0.18 (0.98)	335	0.398	-0.17 (1.00)	335
Different Teacher in M and R	0.375	0.24 (1.03)	333	0.374	0.20 (1.04)	333
Male Teacher	0.493	0.01 (1.01)	331	0.483	-0.17 (1.04)	292
Female Teacher	0.507	-0.01 (1.03)	380	0.517	0.17 (0.93)	385
Urban Area	0.734	0.11 (0.97)	606	0.732	0.11 (0.96)	606
Rural Area	0.266	-0.30 (1.03)	256	0.268	-0.31 (1.04)	256
Public School	0.859	-0.13 (0.92)	686	0.859	-0.10 (0.99)	686
Private School	0.141	0.76 (1.05)	152	0.141	0.54 (0.88)	152
Multigrade School	0.231	-0.28 (1.01)	263	0.234	-0.32 (0.94)	263
Complete School	0.769	0.08 (0.98)	599	0.766	0.10 (1.00)	599

Note: Summary statistics calculated using sampling weights from EN 2004. Standardization of test scores on full sample. The number of observations of collectively exhaustive sub-samples does not always add up to the full sample due to missing information in the sub-group characteristics.

All in all, it is obvious that the sample of same-teacher one-classroom schools is not representative of the national average but rather drawn from the lower part of the knowledge distribution. It is thus questionable if the effect presented in the results section does not only hold for this particular sub-sample. The general effect for the population may be smaller if predominantly worse students profit from teacher subject knowledge, or larger if a minimum amount of student knowledge is necessary as a basis for profiting from teacher subject knowledge.

We address this issue by using a bootstrapping procedure: running regression (2) a large number of times in a sub-sample of the same-teacher one-classroom sample which is comparable to the whole population and extracting the estimated coefficients. This sample (STOC-pop) is constructed by drawing a random sample from the original EN

2004 population that is smaller than the STOC sample and using nearest neighbor propensity score matching to pick those observations from the STOC sample which are comparable to the initial population. The variables along which STOC-pop is made comparable to the initial population are student and teacher test scores in both subjects, and dummies for different school types (urban, public, multigrade). The chosen observations constitute the STOC-pop sample. After running regression (2) on STOC-pop and extracting coefficients and standard errors this procedure is repeated a large number of times.

For a nearest neighbor propensity score matching with caliper 0.01 and 1000 repetitions of the procedure, the regression yields an average teacher test score effect of 0.042 with an average standard error of 0.018 for an average number of 1904 observations. We take this as evidence that the teacher test score effect calculated before is not an artifact of the STOC sample which we draw to identify our regression but that it holds in a similar magnitude in the Peruvian 6th grade student population at large.

Chapter 4

4 The Impact of Natural Hazards on School Progression: Evidence from Rural Peru

4.1 Introduction

One commonly overlooked aspect in the evaluation of natural hazard damages is its effect on educational production. In order to assess the costs of climate change researchers need to account for all of its effects on determinants of economic development; educational production resides prominently among them. Nevertheless, the most influential studies on the impacts of climate change and extreme weather events, the Stern Review on the Economics of Climate Change from 2006 and the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) from 2001[34], do not mention this facet at all. Natural hazards, however, can affect human capital accumulation via a number of channels. Particularly, disasters can distress the health and economic situation of schoolchildren and their families which may result in lower attendance, lower learning and higher dropout; in the long run this will negatively influence economic development.

This paper examines the impact of natural hazards on school progression in rural regions of Peru. The context of this paper is thus a developing country where particularly strong effects of extreme weather events to educational outcomes can be expected (IPCC

[34] Stern (2007), IPCC (2001)

2001). First, developing countries are geographically more exposed to extreme weather events than developed countries. Peru is estimated to be the third most affected country by climate hazards worldwide (Brooks, Adger 2003). Second, developing countries employ large shares of the population in the agricultural sector, which is most exposed to weather conditions. Third, developing countries have lower means of protection and insurance against extreme weather events. Incomplete financial markets and borrowing constraints impede consumption smoothing for many poor people. An extensive literature has found evidence for a worsening of educational outcomes in the face of negative income shocks (see second section).

The contribution of this paper is to extend this more general strand of literature on the effect of negative income shocks on educational outcomes in two respects. First, this paper considers natural hazards in rural regions of Peru as a specific type of negative shock. The effects of natural hazards in developing countries are interesting in their own respect. The third IPCC assessment stressed the vulnerability of poor countries which depend strongly on the primary sector to extreme weather shocks and their dire consequences for poverty reduction. However, we know little about the channels and magnitudes of these effects, which are important, for example to make cost-benefit calculations of climate change mitigation or to specify dynamic general equilibrium models.

Second, most estimates in the literature on the impact of negative income shocks on educational outcomes are vulnerable to the endogeneity of income. This is problematic in the presence of unobserved time-varying variables which may violate the key assumption of strict exogeneity of the regressors. This paper circumvents the endogeneity problem by exploiting natural experiments: I use disaster damage as an exogenous explanatory variable.

The analysis is based on a unique combination of longitudinal data bases at the district level in Peru. I combine information from a national disaster database, an agricultural census and the national school census. The main explanatory variable is the damage created by natural hazards in rural regions measured in hectares of damaged or destroyed farmland. The appeal of this variable is the exogenous nature of disasters which leads to a straight-forward identification strategy in fixed and random effects models.

Also, hectares of farmland is an easily understandable magnitude which can be compared across countries. The key educational outcome examined is the share of enrolled primary schoolchildren being promoted to the next grade in rural areas. Furthermore, I examine the opposite, the share of students failing the grade either due to early withdrawal from school or due to non-promotion by the teacher. This strategy allows distinguishing different channels by which disaster damage causes grade non-completion. Additionally, the analysis is done for boys and girls separately to examine if effects differ depending on gender.

The results of the analysis suggest that damages to farmland caused by natural hazards such as floods, storms and fires have a pronounced impact on school progression in rural Peru. A disaster affecting one hectare of farmland per rural schoolchild decreases the rural district grade completion rate by an estimated 2.4 percentage points on average. Thus, for about 42 hectares of affected farmland one schoolchild is induced to fail the grade he or she is enrolled for. This average, however, understates the effect on small subsistence farmers with only up to 5 or 10 hectares of land. Further analysis suggests that a mean effect of up to one failing schoolchild per 13 hectares of affected farmland is possible if the whole district was owned by farmers with no more than 5 hectares of land each. Applying the mean prediction of estimates to the amount of disasters in Peru, the analysis accounts for about 1500 cases of grade failure in three years. Against common belief, withdrawal from school is not the only channel by which grade failure happens. A significant amount of children remains in school but is not promoted to the next grade – surprisingly, this effect is stronger when subsistence farmland dominates the area. Furthermore, the analysis suggests that there are no statistically significant differences between the effect on boys and girls when disasters occur. Due to large error bounds it is difficult to draw strong conclusions from these additional findings but they certainly represent an interesting area for future research.

Climate change is predicted to produce more frequent and severe weather events thus increasing negative impacts on educational outcomes and long-term human capital formation in developing countries. For example, the Andean Community estimates that Peru will have a 4.4 percent lower GDP with than without climate change by 2025, among other reasons by having 10 percent lower relative agricultural production (CAN

2008). It is imperial to focus further research on this issue to gain a clearer understanding of the channels and magnitudes of weather effects on education.

4.2 Background and Literature Review

Developing countries will be particularly distressed by the consequences of changing climatic circumstances and the increased frequency and severity of extreme weather events. (cf. Stern 2007, IPCC 2001). At the micro level, examples of direct consequences of climate hazards are loss of life, livelihood, private assets and infrastructure. At the macro level this translates into reduced productivity of important economic sectors, especially agriculture. The reasons for these effects are geographic exposure, low incomes, and greater dependence on the most climate sensitive agricultural sector. Latin America in general and Peru in particular are no exception in this respect. Charvériat (2000, p. 94) concludes that "the risk of natural disasters in the Latin American and Caribbean poses a sizable threat to the preservation and continuation of the regional socio-economic development process." Brooks and Adger (2003) from the Tyndall Centre for Climate Change Research conclude that Peru is the third most vulnerable country to climate-related natural disasters worldwide after Honduras and Bangladesh.

As a consequence, extreme weather events severely endanger the livelihood of farmers in developing countries who constitute a considerable share of the total population and are dependent on steady income flows from agriculture. In Latin America, almost 20 percent of the total area are agricultural lands, contributing about 10 percent to the region's GDP and providing occupation for up to 40 percent of the economically active population in some countries (IPPC, Chapter 14). The primary sector remains a key element of regional economies providing work and food security to many of the poorest people in rural regions. Especially subsistence farmers, who only produce for their own consumption, have little means to cope with unexpected shocks to small amounts of farmland as their primary sources of income and will be most affected by extreme weather conditions.

Peru in particular is a country with an accentuated subsistence farming sector which is vulnerable to negative weather shocks. According to the national agricultural census from 1994, 85.3 percent of agricultural production units are in the hand of small landholders with up to 10 hectares of land controlling about 49.4 percent of the total farmland. As a consequence, the largest share of agricultural workers is very exposed to extreme weather events. If a disaster damages farmland of such a small farmer the household is in danger of loosing the majority of income and would have to resort to coping strategies such as child labor. Nevertheless, disaster damage may also affect workers on industrial farms if their employers dismiss parts of the workforce in response to reduced arable surface.

A broader strand of literature has already considered the connection between negative income shocks and educational outcomes in developing countries and has generally found negative impacts. The connection between negative income shocks and decreased schooling runs through an increase in the supply of child labor, especially when there are credit constraints.[35] Such child labor is pervasive in Peru; Patrinos and Psacharopoulos (1997) find that rural children in Peru contribute on average 18 percent to family income. In general, Jacoby and Skoufias (1997) and Beegle, Dehejia, and Gatti (2006), among others, find that households significantly increase market work and decrease school attendance of children in response to anticipated and unanticipated income shocks. Yet, it is always difficult to address concerns of endogeneity of income even in a panel data setting. There may be unobserved time-varying variables correlated with income in different time periods, e.g. through investment decisions, which bias the estimated effect of income shocks.

This work also documents that there are important differences between boys and girls. Overall, girls tend to work more than boys (Edmonds 2007), especially in the household. Domestic work needs to be taken into account when considering the schooling-work trade-off, and can be a primary deterrent to school attendance (cf. Levison and Moe 1998 for Peru). As boys tend to have a smaller work burden in the household than girls, they face fewer barriers to schooling than girls (cf. Assaad et al. 2005 for Egyptian data).

Even though the vulnerability of school attendance to different forms of shock in developing countries is well-established it has so far not entered studies on the analysis of climate change related costs, such as the Stern Review or the IPCC report. Indeed, natural

[35] A great survey of child labor in general and its impact on school attendance is Edmonds (2007).

disasters can negatively affect educational outcomes via a number of channels on the demand and supply side of education. On the supply side, long periods of teaching may be lost due to adverse impacts of extreme weather events on the health and economic circumstances of instructors and the destruction or damaging of schools and relevant other infrastructure. On the demand side, schoolchildren and their families may also be affected in their health and economic situation which may result in lower attendance, lower learning and higher dropout. These channels, especially on the demand side, will be more accentuated in low-income countries. To my knowledge, Holmes (2002) is the only existing study which examines an effect of natural disasters on educational outcomes.[36] He analyses the effect of several hurricanes in North Carolina, USA, on student test score growth using a longitudinal school data set and finds a consistent negative effect.

This paper thus contributes to the thin literature on the connection between weather conditions and educational outcomes in developing countries. First, I use damages caused by extreme weather events as a specific shock which is interesting in its own respect to assess the impact of climate change related events. Second, by the nature of this event, the econometric specification is straight-forward and I can circumvent many potential problems of endogeneity. Third, the impact expressed in hectares of affected farmland is an easily understandable magnitude which can be compared across countries.

4.3 Empirical Implementation

The focus of this paper is on the role of natural hazards in determining grade non-completion. The grade completion probability of a student is a function of two groups of factors: out-of-school factors, such as family and student characteristics, and in-school factors, such as school quality.

A household sending a child to school must have a positive expected value from enrollment. This value is determined by the difference between expected benefits of schooling for the child, such as future wages, and costs, such as fees and opportunity costs of not working in the household or labor market (see for example Gertler, Glewwe 1990).

[36] Porta and Laguna (2007) indicate that hurricane Mitch may be responsible for a reversal of declining drop-out figures in Guatemala which, however, is not substantiated by an analysis.

However, if a poor household is hit by a shock, the expected value from going to school may change for enrolled children who may have to start supporting the family or increase their effort in doing so. This could mean dropping school for the current year or spending less effort on school tasks resulting in higher failure probability.

At an aggregate level, I specify the reduced form model

(1) $y_{dt} = c_1 + \gamma_1 D_{dt} + \mu_d + \delta_t + e_{dt}$

where grade completion rate y in district d at time t is influenced by a shock D which represents disaster damage relative to district size.[37] Also, there are district-level fixed effects μ and time effects δ.

If all students enrolled in school plan on finishing their current grade given current circumstances, e.g. family income, a negative shock to these circumstances may induce some of them to drop out or spend less effort on school, resulting in higher grade non-completion rates in the respective area. In a correctly specified econometric model, I expect to estimate a negative effect of a disaster shock.

The analysis aims to establish whether the hypothesis of non-zero disaster effects can be rejected by estimating the above postulated relationship (1). The variable of interest is shocks to irrigated farmland, measured as hectares of farmland damaged or destroyed by some incidence, such as flooding, drought, fire or similar. In practice, other explanatory variables are not needed in the regression under the assumption that natural hazards are exogenous and thus uncorrelated with factors such as school quality and district poverty.

It is likely that the effect of disaster damage is not uniform across all units but may depend on local circumstances. In particular, a poor subsistence farming unit will be strongly affected by damaged farmland, whereas a large industrial production unit may be less so, especially for small amounts of damage. In (1), γ_1 will thus represent an average effect over all districts independent of their agricultural production structure. It is, however, also reasonable to estimate

[37] With grade completion expressed as a rate, district level variables need to be scaled in order to make them comparable across districts. My approach is to express disaster damage *per student* in the district, i.e., to chose the same denominator as for the dependent variable. The estimated result can thus be expressed as the number of hectares of affected farmland which induce one child less to complete the grade.

$$(2) \qquad y_{dt} = c_2 + \gamma_2 D_{dt} + \gamma_3 A_d + \gamma_4 D_{dt}{*}A_d + \mu_d + \delta_t + \varepsilon_{dt}$$

where disaster damage is interacted with a factor A that characterizes the heterogeneity of districts with respect to their vulnerability to disaster shocks. In practice, I express this factor as the share of farmland in the district worked on by small-scale subsistence farmers.

It remains to be determined if (1) and (2) should be estimated by fixed effects or random effects regression. This leads to the question whether the observed shocks are truly uncorrelated with unobserved district fixed effects. If natural shocks are random only conditionally on an unobserved propensity for shocks, this problem can be solved via fixed effects regression.[38] Through demeaning of equations at several points in time, the problematic fixed effects including regional propensity for shocks are removed, such that there is no remaining correlation of the incidence of shocks and unobserved time-constant factors. If natural shocks are unconditionally random, both fixed and random effects estimation will be consistent but random effects will be efficient. The Hausman test can help to determine which specification to use.

A remaining source of concern is the possibility of unobserved time-varying factors which violate the assumption of strict exogeneity of the shock variables. This could be the case if shocks have consequences which last for more than one period. For example, shocks may hit a community so hard that the poorest and possibly those students with lowest propensity to complete the grade drop out of school permanently. The unobserved overall ability of students may thus be higher in the next period, creating a correlation of ability and deviation from the mean shock.

[38] Some areas are more likely to be hit by a shock than others, such as earthquake-prone areas, mountainous regions for droughts or wetlands for floods. If the resulting human conglomerations in this region evolve dependent on the area's specificities then the area's propensity for shocks is not exogenous in a cross-sectional regression. It may be correlated first of all with income and through this channel with other influential factors, such as vulnerability to shocks, school quality and individual ability, motivation and opportunity cost of schooling.

4.4 Data

Different data sources are merged at the district level in order to enable estimation: data on 1) natural hazards, 2) on the distribution of farmland ownership and 3) on school completion and failure rates.

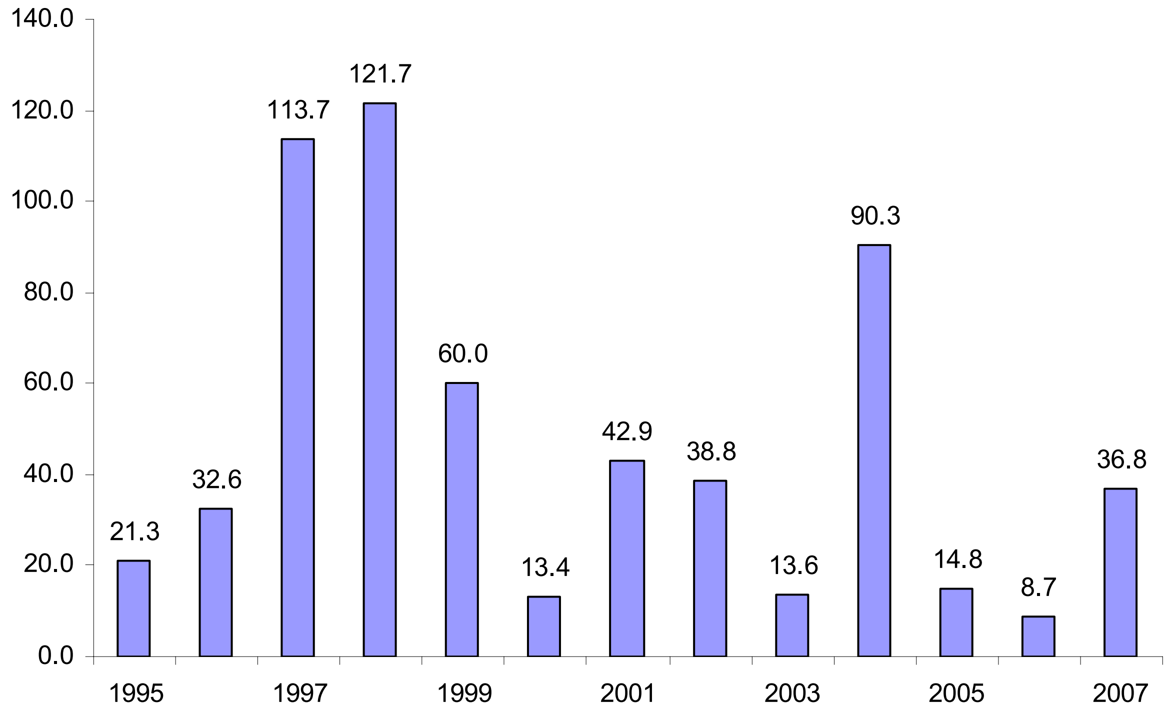
4.4.1 Natural Disaster Data

Information on natural hazards is retrieved from the Peruvian internet platform SINPAD (Sistema Nacional de Información para la Prevención y Atención de Desastres) which is part of the larger National Civil Defense System run by the National Civil Defense Institute. It is an internet-reporting system of damages and needs due to natural and human-caused disasters. The tool is used for official disaster monitoring, prevention and reaction. It is a uniquely comprehensive database which permits registering every single disaster in the country since 1995. For example, in the period January 1st 2003 to December 31st 2005 it contained more than 12,000 incidences of disaster in Peru.

Disasters contained in SINPAD are classified by phenomenon and linked to their respective damage, both in detailed categories. The types of phenomena contained are geodynamic, meteorological, biological or technological disasters. Also, the exact phenomenon is detailed, e.g., volcanic activity, earthquake, flooding, storm, landslide, drought, or fire. Furthermore, SINPAD registers the damages of disasters. These are classified into damages to life and health, buildings, ways of transport, agricultural infrastructure, farmland, crops and animals. In addition, the exact quantified damage is listed in the database, e.g., number of people deceased, hurt, affected, houses damaged and destroyed, or streets damaged and destroyed.

This paper uses information on damages to farmland as the main explanatory variable to measure shocks to rural areas and households. Within the disaster quantifications, SINPAD reports hectares of damaged and destroyed farmland which I aggregate to *affected farmland* as my main explanatory variable. These damages are clearly confined to rural areas. Figure 4-1 shows the pattern of yearly reported hectares of destroyed farmland since 1995 and demonstrates that destruction of farmland by natural hazards is not a stable process. Instead, destroyed farmland varies between less than 10,000 and more than 100,000 annual hectares.

Figure 4-1. Destroyed Farmland in 1000 hectares, 1995-2007



Source: SINPAD.

Table 4-1 shows the frequency and severity of disasters which affect farmland. Droughts are by far the most frequent and severe force in damaging and destroying farmland with 29 percent of incidences and 60 percent of affected farmland. Other high-frequency disaster phenomena with more than twenty incidences per year are frost, flooding, rain, hail and storm. Together, they account for more than 90 percent of affected farmland. Landslides, infestations and high tides are also on the list of severe disasters with more than 100 affected hectares per incident.

Table 4-1. Disasters Affecting Farmland, All Districts, January-December, 2003-2005

Phenomenon	Freq.	Percent Incidences	Affected Farmland (ha)	Percent Farmland Affected	Affected Farmland / Incident
Drought	303	29.2	339,234	60.45	1119.6
Frost	204	19.7	109,931	19.59	538.9
Flooding	172	16.6	58,643	10.45	340.9
Rain	103	9.9	19,088	3.40	185.3
Hail	74	7.1	7,778	1.39	105.1
Storm	60	5.8	7,203	1.28	120.1
Landslide	29	2.8	2,512	0.45	86.6
Fire	20	1.9	10,173	1.81	508.7
Flash Flood	20	1.9	680	0.12	34.0
Snow	14	1.4	681	0.12	48.6
High Tide	13	1.3	4,013	0.72	308.7
Alluvion	10	1.0	508	0.09	50.8
Collapse	6	0.6	301	0.05	50.2
Others (Ext. Geodynamic)	3	0.3	6	0.00	2.0
Others (Meteorologic)	2	0.2	57	0.01	28.5
Others (Int. Geodynamic)	2	0.2	14	0.00	7.0
Infestation	1	0.1	380	0.07	380.0
Earthquake	1	0.1	1	0.00	1.0
Total	1,037	100	561,203	100	541.2

Source: own calculations based on SINPAD.

The database dates back to 1995. While early years may suffer from underreporting there is no reason to believe that this is the cause for recent years. Disaster reports which end up in SINPAD are filed by local civil defense committees which exist in all districts in Peru. The filing of disaster reports is linked to the reception of aid measures which are also contained in the database. There is thus no reason to believe that reporting is endogenous, e.g., that poorer districts may not report all of their disasters.

Table 4-2 summarizes the frequency and severity of disasters affecting farmland that are relevant for the analysis. The sample of included districts and affected farmland figures is reduced in two ways compared to the whole population of disasters: excluding affected farmland in the months January to March and excluding districts which during those three years at some point of time reported a drought.

Hectares of farmland damaged or destroyed are added up from April to December of the respective year because this is the period of time which the school year spans. As a result, schoolchildren enrolled at the beginning of the school year in April will be affected by shocks to farmland from April on.

Table 4-2. Summary Statistics for Affected Farmland April-December, Excluding Drought-districts

Affected Farmland (Ha)	2003	2004	2005
All Districts			
Observations	1407	1407	1407
Mean	6.8	9.2	5.8
S.D.	73.5	98.3	138.1
Affected Districts			
Observations	49	64	32
Mean	193.8	202.2	252.8
S.D.	348.1	419.8	894.8
Max	1762	2374	5000
Total	9498	12939	8091
Affected Districts *(Ha/rural student)*			
Observations	49	64	32
Mean	0.14	0.19	0.23
S.D.	0.24	0.50	0.71
Max	1.05	3.43	3.82

Source: own calculations based on SINPAD.

Drought-affected districts are excluded from analysis because droughts are longer term events and do not hit districts by surprise. For example, the department Tacna was reportedly in state of drought for more than two years during 2003 to 2005 while the incidence of drought was only reported once in the disaster database much later than the actual onset. The date of reporting thus cannot correspond to a day-specific realization of the disaster and cannot be congruently classified as before or after the start of the school year. This inaccuracy leads to the following problem: in case of reporting during the school year while the onset of the drought was before April, the drought will already have influenced the enrollment decision, an effect which I am not able to measure. Also, for droughts lasting longer than one year, the strict exogeneity assumption is less likely to be fulfilled. As a consequence, I find it most reasonable to exclude drought affected districts altogether even though they account by far for most of the damaged and destroyed farmland. However, it seems reasonable to believe that the estimated impact should be valid for all affected farmland independent of the disaster type which caused it, including droughts.

The upper panel of Table 4-2 contains the disaster statistics measured in hectares of affected farmland for all districts including those which were not affected. The middle and lower panel show the disaster statistics only for affected districts – while the middle panel shows them in hectares of affected farmland the lower panel scales the statistics relative to the district primary school population, which is the relative measure later used in the regressions in order to make districts comparable in size.

During 2003-2005, out of 1662 districts 1407 were never affected by a drought. In 2003, 49 non-drought districts registered disaster-affected farmland, in 2004 64 districts, and in 2005 32 districts. Overall, 121 non-drought districts were at least once affected by a disaster-caused destruction or damaging of farmland.

Note that there are enormous differences between the sum of disaster affected farmland in Table 4-2 and Table 4-3. The difference stems from two sources: first, most disasters in Peru happen during the months of January to March; about 90 percent of farmland was affected during those months, as we can see in Table 4-3. Second, drought-affected districts are excluded from the regression sample. As a consequence, not only hectares of farmland damaged or destroyed by droughts are lost for the analysis but also all the remaining disasters in all three years of drought-affected districts.

Table 4-3. Total of Disaster-Affected Farmland, by Time of Year

Shocks to Farmland (Ha)	January - March	April - December
2003	40,975	9,772
2004	274,186	29,264
2005	183,678	23,328
Total	498,839	62,364

Source: own calculations based on SINPAD.

4.4.2 Agricultural Production Structure Data

In order to account for the structure of agricultural production in Peruvian districts, I use the Peruvian agricultural census (CENAGRO) from 1994 to approximate the share of land held by subsistence farmers in affected districts 2003 to 2005.

The CENAGRO was carried out between October and November of 1994 by the National Institute of Statistics and Informatics (INEI) and covered the entire country, excluding totally urbanized areas. The focal statistical unit is the agricultural unit, defined as any piece of land consisting of one or more parcels, totally or partially used for agricultural production, carried out as an economic unit by the agricultural holder, without regard to size, tenure or legal status. Data on holding and holder characteristics, tenure, land use and livestock etc. were collected through direct interview. The data used for this analysis only refers to agriculturally used land, not total surface.

Table 4-4. Total of Disaster-Affected Farmland, by Time of Year

Farmland production structure	National	Affected Districts	Affected Non-drought
Districts	1662	177	121
All sizes			
No. of units (in 1000s)	1671.22	234.15	170.72
Surface (in 1000 has)	5476.98	970.77	771.55
Average surface / unit (in has)	3.28	4.15	4.52
Units with <10 ha farmland			
Share of units	0.853	0.756	0.704
Share of surface held	0.494	0.407	0.362
Units with <5 ha farmland			
Share of units	0.710	0.579	0.511
Share of surface held	0.299	0.221	0.186

Source: own calculations based on CENAGRO.

Table 4-4 displays an overview of agricultural units in 1994, nationally compared to disaster-affected districts (in 2003 to 2005). Peru has almost 1.7 million agricultural units holding almost 5.5 million hectares of farmland. The average size is 3.3 hectares, compared to 4.2 hectares for all disaster-affected districts and 4.5 for disaster-affected non-drought districts. Nationally, about 85 percent of units work on less than 10 hectares, making up about 49 percent of the total farmland. In affected districts, 76 percent of units have less than 10 hectares and constitute 41 percent of total farmland. Typical subsistence farmers without production for market survive on 1 to 5 hectares of land (Plaza, Stromquist 2006) constituting 71 percent of units nationally and 30 percent of total surface. In affected districts, 58 percent of units work on less than 5 hectares, constituting 22 percent of total surface. Disaster-affected districts seem to be slightly bigger and less small-scale than the national average.

The numbers reveal that much of agricultural production in Peru is carried out on a very small scale, with the grand majority of farms working on less than 10 or 5 hectares of land. These farms and their workers can largely be considered subsistence farmers only producing for their own consumption. These production units are particularly exposed to extreme weather shocks due to lack of coping strategies.

4.4.3 Grade Completion Data

The third database used is the Peruvian school census 2004, 2005 and 2006 which covers all Peruvian educational institutions. It is collected on a yearly basis by the Ministry of Education via questionnaires specific to the type and level of institution. The survey information is self-reported by the schools. Schools are identified via a unique identification code and can thus be followed over time. Information is collected to reflect school registers at the date of May 30. The information collected does not allow for individual student profiles but aggregation at the grade and school level. For example, information contains the grade structure of students according to gender, age, native language and repeater status but it is not possible to follow who exactly is failing the grade.

Information on grade completion is collected for the previous school year such that the resulting database covers the end of year results for the years 2003, 2004 and 2005. The big advantage of the school census data is that it can be used to compare beginning of year enrollment with end of year results. Enrollment at the beginning of the year constitutes a revealed preference and signals positive utility from schooling. The panel setting with random effects allows using non-disaster affected districts as a control group to identify the average year specific rate of non-completion in the absence of disaster treatment.

In the school database, the enrollment and completion statistics are cleaned and added up to district figures only for schools in rural and marginal urban areas. Schools with inconsistent reporting or large enrollment changes over the years are excluded.[39] The subset used for estimation is a cleaned sample of formal, non-adult primary schools which are observed in all three years.

Table 4-5 summarizes the rural student population by year. In 2003, 81.5 percent of all rural primary schoolchildren completed the grade meaning that almost one in five children failed the grade. About 10 percent of children failed due to non-promotion while about 8 percent failed due to early withdrawal or low attendance. There is a small but steady upward trend in promotion rates between 2003 and 2005.

Table 4-5. Summary Statistics Rural Student Population

	2003		2004		2005	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Share Boys	0.511	(0.026)	0.511	(0.026)	0.510	(0.026)
All						
Share Promoted	0.815	(0.074)	0.827	(0.071)	0.833	(0.071)
Share Not Promoted	0.102	(0.043)	0.097	(0.045)	0.093	(0.044)
Share Withdrawn	0.082	(0.041)	0.076	(0.036)	0.073	(0.036)
Boys						
Share Promoted	0.815	(0.071)	0.825	(0.068)	0.831	(0.068)
Share Not Promoted	0.104	(0.043)	0.098	(0.044)	0.096	(0.043)
Share Withdrawn	0.081	(0.040)	0.076	(0.035)	0.073	(0.036)
Girls						
Share Promoted	0.816	(0.081)	0.829	(0.077)	0.835	(0.077)
Share Not Promoted	0.101	(0.047)	0.095	(0.048)	0.091	(0.047)
Share Withdrawn	0.082	(0.044)	0.076	(0.039)	0.073	(0.039)

Source: own calculations based on national school census, 2004-2006.

[39] Large changes in enrollment between years can stem either from structural breaks or strong measurement error, bot of which should be avoided. For example, in a school losing a large share of students due to a natural hazard in the area the remaining children may reflect a very different socioeconomic and ability structure and level. The school will thus not be comparable anymore and the strict exogeneity assumption needed for consistent estimation is violated. As a consequence, I may underestimate the effect of natural hazards if schools which are most severely hit are dropped from the sample.

The table indicates that there is no gender discrimination in rural Peru which can be found in the school system of many other developing countries. First, boys and girls constitute almost equal shares in the student population. Second, completion, non-promotion and withdrawal rates are equal among boys and girls in all years.

4.5 Results

Table 4-6 documents the main results of the paper from fixed (FE) and random effects (RE) estimation of specification (1), with and without year effects. The Breusch and Pagan Lagrangian multiplier test (not shown in table) strongly suggest the existence of individual effects such that FE and RE models are appropriate. The dependent variable is the district-level share of students successfully completing the school year and being promoted to the next grade. The independent variable of interest is the district-level amount of disaster-affected (damaged plus destroyed) farmland in hectares per rural student. The scaling is done to make bigger and smaller districts comparable. The estimation is performed on all districts which had not registered any drought throughout the period 2003 – 2005. The Hausman test strongly rejects significant differences between results from the fixed and random effects specification. This supports the view that natural disasters are unconditionally random events. In the following discussion I thus only report random effects.

Table 4-6. Disaster Impact on Grade Completion for Fixed and Random Effects Estimation

Dep. Variable: Share Completed	FE	RE	FE	RE
Farmland affected / rural schoolchild	-0.0220**	-0.0233**	-0.0225***	-0.0237***
	(0.0086)	(0.0093)	(0.0085)	(0.0087)
Constant	0.8085***	0.8084***	0.7970***	0.7969***
	(0.0001)	(0.0020)	(0.0010)	(0.0023)
Year Dummies	No	No	Yes	Yes
R^2	0.001	0.001	0.012	0.012
F (FE) / Chi2 (RE)	6.55	6.28	48.72	160.50
Observations	4221	4221	4221	4221
Hausman Test p-value	0.57		0.95	

Note: Robust standard errors in brackets. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. Hausman test computed on regressions without robust standard errors which produce standard errors that never change the significance level of significant coefficients by more than 0.005 as compared to robust standard errors.

Both fixed and random effects results confirm a significant negative effect of disaster affected farmland on district grade promotion shares in rural primary schools. As the inclusion of year effects does not change the results (columns 3 and 4) we can be more confident about the exogenous nature of the explanatory variable. According to the results, a natural hazard damaging or destroying one hectare of farmland per rural schoolchild reduces the share of promoted students by 2.2 to 2.4 percentage points. In other words, for about 42 hectares of affected land, one schoolchild is not promoted to the next grade. Considering the total quantity of disaster-affected farmland between April and December 2003 to 2005, 62,364 hectares (Table 4-3), disasters forced about 1,500 students to fail the grade according to the mean prediction, not including those who were deterred from enrolling. If we assume that the same effect magnitude holds for children that were deterred from enrolling by disaster damage of about 500,000 hectares between January and March during those three years, almost 12,000 children were prevented to enroll in school.

Table 4-7 shows the results for random effects estimations according to specification (2), i.e., including affected farmland, a baseline effect for the share of small-scale agricultural units in the district, and an interaction term between the two. In the left panel, small-scale agricultural units are defined as those with less than 10 hectares of farmland, in the right panel, those with less than 5 hectares of farmland. In each panel, the first column contains only the non-interacted variables, the second column adds the interaction effect, and the third column assumes that there is no average disaster effect independent of the interaction effect.

Table 4-7. Disaster Impact on Grade Completion for Random Effects Estimation

	Subsistence farmland: < 10 has			Subsistence farmland: < 5 has		
Dep. Variable: Share Completed	[1]	[2]	[3]	[4]	[5]	[6]
Farmland affected / rural schoolchild	-0.0218**	-0.0165		-0.0223***	-0.0190*	
	(0.0086)	(0.0155)		(0.0086)	(0.0113)	
Share of subsistence farmland	-0.0471***	-0.0466***	-0.0460***	-0.0568***	-0.0557***	-0.0548***
	(0.0122)	(0.0122)	(0.0122)	(0.0180)	(0.0182)	(0.0182)
Farmland affected / rural schoolchild		-0.0176	-0.0546**		-0.0238	-0.0755**
x Share of subsistence farmland		(0.0426)	(0.0232)		(0.0506)	(0.0363)
Constant	0.7988***	0.7987***	0.7987***	0.7983***	0.7982***	0.7982***
	(0.0023)	(0.0023)	(0.0023)	(0.0023)	(0.0023)	(0.0023)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.019	0.020	0.019	0.017	0.017	0.017
Chi2	175.50	175.70	174.70	170.20	170.80	168.70
Observations	4221	4221	4221	4221	4221	4221

Note: Robust standard errors in brackets. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

For both definitions of small-scale farms, the results of column 1 and 4 do not change markedly compared to Table 4-6 in the presence of subsistence farmland share as an additional regressor. Interestingly, the share of subsistence farmland is significantly associated with lower grade completion rates, possibly due to correlation with district-wide effects such as higher poverty or lower school quality. In columns 2 and 4, which include the interaction effect, affected farmland looses in magnitude and significance, dropping to -0.017 and -0.019. The interaction effect turns out negative for both subsistence definitions (-0.018 and -0.024) but insignificant, likely due to high correlation (0.99 and 0.97) with non-interacted farmland. However, in a Wald test (not shown), farmland and the interaction term are jointly significant at the 10 (colum 2) and 5 (column 4) percent level. Assuming that the mean effect is consistently estimated there would be both an average effect of disaster affected farmland and an effect that depends on the structure of farm holdings in the district. If every farm in the district is less than 10 hectares in size, the aggregate effect on grade completion is -0.033, for all farms having less than 5 hectares it is -0.043. In case of 100 percent subsistence farmland, the aggregate effect is thus such that 29 or 23 hectares of farmland devastation would cause one schoolchild to fail the grade.

Columns 3 and 6 of Table 4-7 confirm that there is a stronger negative effect of disasters if more farmland is held by smallholders, assuming that there is no average effect but only one that is dependent on local farmland structure. Column 3 shows a significant negative effect of -0.055 of the interaction term between affected farmland and the share of farms smaller than 10 hectares. According to RE, one child is induced to fail the grade for every 18 hectares of damage if farmland is completely held by agricultural units of less than 10 hectares. Column 6 reveals an even bigger significant effect of -0.076 such that one child is induced to fail the grade for every 13 hectares if farmland is completely used by agricultural units of less than 5 hectares.

Table 4-8 shows the results for random effects estimations similar to the one before only with the share of students not promoted and the share of students withdrawing from school before the end of the school year as separate dependent variables. These two categories constitute the aggregate of students failing the grade, i.e., 100 percent minus the completion share.[40] While we are ultimately interested in the share of those failing the grade due to the disaster for whatever reason an analysis of these categories separately may still be useful to learn more about the consequences of disasters.

Table 4-8. Disaster Impact on Grade Failure from Random Effects Estimation

			Subsistence farmland: < 10 hectares		Subsistence farmland: < 5 hectares	
Dep. Variable: Share	Not promoted	Withdrawn	Not promoted	Withdrawn	Not promoted	Withdrawn
Farmland affected / rural schoolchild	0.0078	0.0164*	-0.0154*	0.0314**	-0.0067	0.0255**
	(0.0071)	(0.0099)	(0.0093)	(0.0158)	(0.0047)	(0.0126)
Share of subsistence farmland			0.0425***	0.004	0.0523***	0.0033
			(0.0086)	(0.0067)	(0.0131)	(0.0099)
Farmland affected / rural schoolchild x Share of subsistence farmland			0.0703*	-0.0512*	0.0907**	-0.0649*
			(0.0383)	(0.0280)	(0.0439)	(0.0333)
Constant	0.1076***	0.0952***	0.1060***	0.0950***	0.1064***	0.0951***
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.006	0.009	0.022	0.010	0.019	0.010
Chi2	62.06	97.57	91.03	98.77	84.10	99.17
Observations	4221	4221	4221	4221	4221	4221

Note: Robust standard errors in brackets. Significance levels: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

The point estimates for farmland affected by disaster per rural student indicate that about two thirds of those students who fail the grade because of a disaster shock do so because of withdrawal from school while about one third remains in school but is not promoted at the end of the year. Yet, the error bounds are too large to distinguish the

[40] A marginal difference between the sum of non-promotion and withdrawal share compared to 100 percent minus the promotion share can arise from students deceasing throughout the year who do not belong to either category.

effect sizes with certainty. When looking at the middle and right panels with a subsistence farmland effect plus its interaction, a different pattern emerges: it seems as if subsistence farming is much more associated with non-promotion than with withdrawal. The share of subsistence farmland is in both cases strongly associated with higher non-promotion but not higher withdrawal shares (second row). The average effect of damaged farmland is negative on the non-promotion share and positive on the withdrawal share (first row). In contrast, the interaction effect expresses that, compared to the average effect, a disaster in subsistence farmland areas drives up the non-promotion share and decreases the withdrawal share (third row). It is, however, unclear if the effects can be distinguished statistically. One possible reason for this pattern is that the rules for failing students due to insufficient attendance may be more relaxed in areas where many children frequently have to work in the fields. Another possibility is that workers in poorer subsistence farm areas may be more inclined to leave their children in school without intention to complete, e.g., for the benefit of free school meals.

Table 4-9 shows the results from random effects estimation on the sub-samples of boys and girls with the completion share of students as the dependent variable. In both cases, there is a negative effect of disaster affected farmland on the within-gender promotion share. This effect is only significant for boys, and the point estimate for boys (-0.030) is stronger than for girls (-0.020). Similarly, the disaster effect is stronger for boys than girls for the interaction with the share of subsistence farmland. However, the error bounds of these estimates are large such that equal effects for boys and girls cannot be rejected.

Table 4-9. Disaster Impact on Grade Completion for Boys and Girls from Random Effects Estimation

	Boys			Girls		
Dep. Variable: Share Completed	[1]	[2]	[3]	[4]	[5]	[6]
Farmland affected / rural schoolchild	-0.0299***	-0.0212***		-0.0200	-0.0214	
	(0.0075)	(0.0080)		(0.0146)	(0.0197)	
Share of subsistence farmland (<5 has)	-0.0476***	-0.0448**	-0.0437**	-0.0665***	-0.0669***	-0.0659***
	(0.0179)	(0.0178)	(0.0178)	(0.0187)	(0.0190)	(0.0190)
Farmland affected / rural schoolchild		-0.0618	-0.1195*		0.0102	-0.0480
x Share of subsistence farmland (<5 has)		(0.0768)	(0.0620)		(0.0536)	(0.0300)
Constant	0.7980***	0.7979***	0.7978***	0.7984***	0.7984***	0.7983***
	(0.0024)	(0.0024)	(0.0024)	(0.0026)	(0.0026)	(0.0026)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.014	0.015	0.014	0.016	0.016	0.016
Chi2	117.50	126.90	107.50	137.70	138.80	140.20
Observations	4218	4218	4218	4221	4221	4221

Note: Robust standard errors in brackets. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 4-10 displays the results from random effects estimation on the share of non-promotion and withdrawal for boys and girls, for farmland affected and separately for affected farmland interacted with the two shares of smallholdings. For boys, all estimates suggest that the majority of grade failure due to disaster damage comes from non-promotion, not withdrawal. For girls, the picture is more ambiguous: the estimate for average affected farmland suggests that the effect for girls comes through withdrawal. When including the interaction effect, however, the pattern is the same as before: disaster shocks on subsistence farmland increase non-promotion, not withdrawal.

Table 4-10. Disaster Impact on Grade Failure for Boys and Girls from Random Effects Estimation

	Boys		Girls	
Dep. Var.: Share	Not Promoted	With-drawn	Not Promoted	With-drawn
A. Farmland affected / rural schoolchild	0.0216	0.0084	-0.0027	0.0226**
	(0.0132)	(0.0118)	(0.0088)	(0.0114)
B. Farmland affected / rural schoolchild	0.0091	0.0118	-0.0159*	0.0370***
	(0.0160)	(0.0158)	(0.0091)	(0.0133)
Farmland affected / rural schoolchild	0.0897	-0.0241	0.0939**	-0.1026**
x Share of subsistence farmland (<5ha)	(0.0640)	(0.0408)	(0.0394)	(0.0410)
Observations	4218	4218	4221	4221

Note: Results from random effects regression including the share of subsistence farmland and year dummies. Robust standard errors in brackets.

Overall, the results suggest that withdrawal from school, and possibly child labor, is not the only channel through which economic shocks hinder the school progression of rural children. Especially in areas characterized by subsistence farmland, disasters do not drive children out of school but lead to higher non-promotion shares. One possible explanation is that disasters induce a shift of time allocation from learning and homework to actual work in the labor market or tasks in the household which prevent sufficient learning achievement for grade promotion. As a result, the main counteracting policy could be both financial relief for households to deter child work but also improvements in the curriculum to transmit the learning materials more effectively in the available time.

4.6 Conclusion

This essay is one of the first to address the detrimental effects of climate change on human capital formation by estimating the effect of natural hazard damage in hectares of farmland on grade completion rates in rural Peru. I find that there is a significant detrimental impact of natural disasters on school progression. On average, 42 hectares of affected farmland are predicted to cause on student not to complete the grade. The estimate thus accounts for a mean prediction of about 1500 failing students in three years. The frequency and severity of climatic events is predicted to increase in the future thus creating more damage to human capital formation.

The focus of this paper is narrow as it only considers grade completion effects of farmland which is destroyed by catastrophic events. We can imagine many more channels of climatic events to affect the supply and demand of education, especially preventing children to constantly attend school and thus disrupting learning. Examples are the destruction of transportation and schooling infrastructure or health effects on teachers and students. Economic effects do not only stem from disaster damage. Maybe more importantly, temperature changes affect the productivity of farmland. Also, climate change is predicted to spur migration movements in developing countries which will certainly also reduce the amount of schooling attainable for children. Certainly, there is much more to learn about these effects to design appropriate coping strategies and prevent the loss of human capital which is so crucial for developing countries.

References

Assaad, Ragui, Levison, Deborah, Zibani, Nadia (2005). *The Effect of Child Work on Schooling: Evidence from Egypt.* Unpublished manuscript.

Beegle, Kathleen, Dehejia, Rajeev H., Gatti, Roberta (2003). Child Labor, Income Shocks, and Access to Credit. World Bank Policy Research Working Paper No. 3075.

Brooks, Nick, Adger, Neil (2003). Country Level Risk Measures of Climate-related Natural Disasters and Implications for Adaptation to Climate Change. Tyndall Centre Working Paper No. 26. Norwich.

CAN (2008). *El Cambio Climático no tiene fronteras. Impacto del Cambio Climático en la Comunidad Andina.* Lima: La Comunidad Andina.

Charvériat, Céline (2000). Natural Disasters in Latin America and the Caribbean: An Overview of Risk. Inter-American Development Bank Working Paper No. 434. Washington, D.C.

Edmonds, Eric V. (2007). Chapter 57 Child Labor. In T. P. Schultz & John A. Strauss (Eds.), *Handbook of Development Economics.* Volume 4, pp. 3607–3709. Elsevier.

Gertler, Paul, Glewwe, Paul (1990). The Willingness to Pay for Education in Developing Countries: Evidence from Rural Peru. *Journal of Public Economics, 42*(3), pp. 251–275.

Holmes, George M. (2002). Effect of Extreme Weather Events on Student Test Performance. *Natural Hazards Review, 3*(3), pp. 82–91.

IPCC (2001). *Climate Change 2001: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge: Cambridge Univ. Press.

Jacoby, Hanan G., Skoufias, Emmanuel (1997). Risk, Financial Markets, and Human Capital in a Developing Country. *Review of Economic Studies, 64*(3), pp. 311–335.

Levison, Deborah, Moe, Karine S. (1998). Household Work as a Deterrent to Schooling: An Analysis of Adolescent Girls in Peru. *Journal of Developing Areas, 32*(3), pp. 339–355.

Patrinos, Harry A., Psacharopoulos, George (1997). Family Size, Schooling and Child Labor in Peru – An Empirical Analysis. *Journal of Population Economics, 10*(4), pp. 387–405.

Plaza, Orlando, Stromquist, Nelly P. (2006). Consequences of Structural Adjustment on Economic and Social Domains: Two Decades in the Life of Peru. *The ANNALS of the American Academy of Political and Social Science, 606*(1), pp. 95–115.

Porta, Emilio, Laguna, José R. (2007). Equidad de la Educación en Guatemala [Education Equity in Guatemala]. No. 4. Guatemala City: USAID Guatemala; Academy for Educational Development.

Stern, Nicholas (2007). *The Economics of Climate Change: The Stern Review.* Cambridge: Cambridge Univ. Press.

Chapter 5

5 Connecting the Unobserved Dots: A Decomposition Analysis of Earnings Inequality Changes in Urban Argentina, 1980-2002

5.1 Introduction

Earnings inequality in Argentina decreased during the 1980s and then rose during the 1990s.[41] During the 1990s, the rise in inequality was partially driven by increases in the returns to education. Candidate explanations for the increase in the returns to education include trade-related changes in production, changes in institutions like the minimum wage and union membership, and skill-biased technical change (SBTC). Available evidence suggests that the first two candidate explanations do not appear to have driven changes in returns to education, leaving skill-biased technical change as the most likely cause. We examine the Argentina case for changes in the returns to unobserved skill by decomposing the variance of earnings over time. The decomposition shows that the residual wage variance increased in Argentina in the 1990s. Under plausible assumptions, this implies that the returns to unobserved skill have risen. Following on similar analysis done for the United States, we interpret the increase in the returns to unobserved skill as evidence for skill-biased technical change.

[41] This chapter is based on Demombynes and Metzler (2008).

To decompose changes in the variance of earnings, we employ Lemieux's (2002) semi-parametric re-weighting technique. The method accounts for a change in the composition of observable workers' characteristics over time, allowing us to estimate a counterfactual wage distribution which holds initial population characteristics constant over time. The change in the overall variance of wages can be decomposed into changes due to changes in observed skills, changes in the returns to observed skills, and changes in the returns to unobserved skill.

The change in the returns to unobserved skill also has the potential to explain a puzzle regarding informal labor in Argentina. Argentina has seen a long-run shift from formal to informal wage employment. During the 1990s rates of informal salaried employment increased while at the same the wages of informal workers, relative to formal workers fell. The growing gap in wages between formal and informal workers is not explained by the rising returns to education. The growing gap may be driven, however, by changes in the demand for unobserved skills driven by SBTC. If employers observe these skills, there will be less demand for workers without them. Skilled workers may thus be hired into formal jobs, which would consistently explain the increase in informality as well as the increasing informal-formal wage gap.

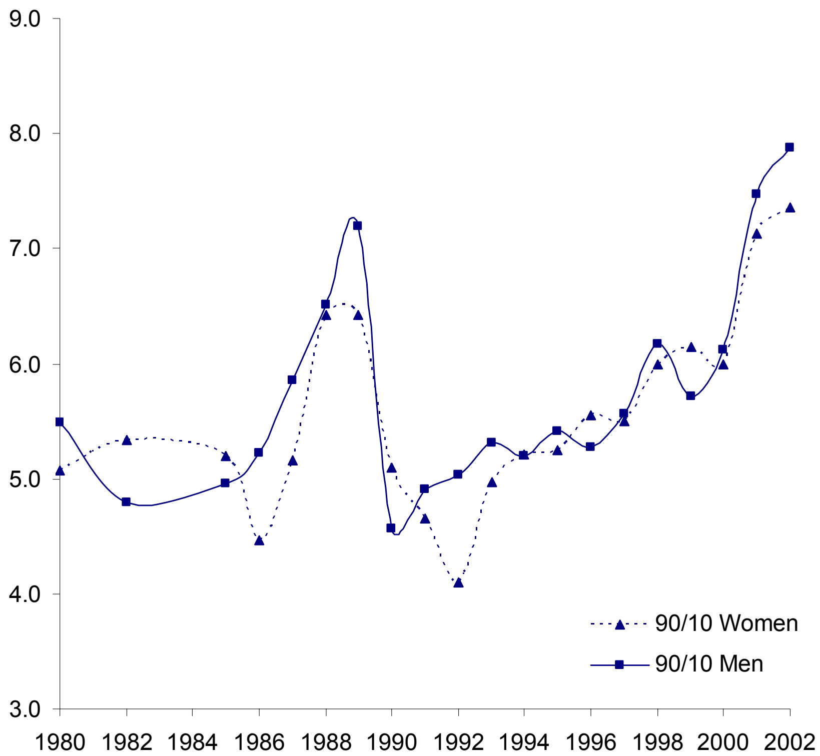
We find that the returns to unobserved skill have increased particularly between the median and the bottom of the distribution, where informal salaried workers are concentrated. This is compatible with the idea that an increase in the returns to unobserved skill has been responsible for the increased gap between formal and informal wages.

5.2 Background

5.2.1 Earnings Inequality and the Distribution and Remuneration of Skills in Argentina

Argentina has seen pronounced economic cycles in recent decades. The 1980s were characterized by mostly weak economic performance in Argentina. An economic crisis in 1981-82 was followed by a short recovery, and then a new crisis in 1985. In the late 1980s, Argentina was again marked by an economic crisis and then hyperinflation. A new stabilization plan was initiated and a fixed exchange rate was implemented in April 1991. The currency board was paired with a strategy of trade liberalization, deregulation and privatization, resulting in a stable economy for most of the 1990s. Signs of renewed crisis were felt by the end of the 1990s when the Argentina's public debt had mounted to unsustainable heights. International financial markets put huge pressure onto the Argentine currency, culminating in the abolishment of the currency board in January 2002. Since 2003, the Argentine economy has been recovering.

Figure 5-1. Earnings Inequality in Gran Buenos Aires: 90:10 Ratio, 1980-2002



Note: Considered are hourly wages for full-time workers above age 15 with one occupation.

The 1980s and the 1990s show very different patterns of earnings inequality changes (see Figure 5-1). The first half of the 1980s in Argentina was characterized by falling earnings inequality. Between 1980 and 1986 the ratio of the 90th to the 10th percentile of the earnings distribution fell from 5.5 to 5.2 for men and from 5.1 to 4.5 for women. Following the late 80s period of high inflation, during which measured wage inequality jumped temporarily due to increased measurement error, wage inequality in

1992 was slightly lower than in 1986. Then, between 1992 and 2002, wage inequality increased sharply; the 90:10 ratio jumped from 5.0 to 7.9 for men and from 4.1 to 7.4 for women.

The changes in earnings inequality are in part driven by changes in the levels and skills and their returns.[42] As Table 5-1 shows, the returns to higher education decreased during the 1980s and increased in the 1990s.[43] The returns of completed high school education compared to primary education for men were 0.64 in 1980, dropping 0.45 in 1992, and rising to 0.56 in 2002. For women, the marginal returns of high school education were 0.59 in 1980, 0.37 in 1992, and 0.72 in 2002.

Table 5-1. Marginal Returns to Education Levels

	Marginal Effect of Education			
	MEN		WOMEN	
	Completed High School	Completed College	Completed High School	Completed College
1980	0.64	0.59	0.59	0.53
1986	0.52	0.70	0.54	0.45
1992	0.45	0.66	0.37	0.38
1998	0.47	0.98	0.66	0.69
2002	0.56	1.03	0.72	0.79

Note: Marginal returns to education are estimated by specifying standard Mincer wage regressions, in which the dependent variable is log hourly wages from principal occupations for adult full-time workers with one single job. Regressors include a quadratic in potential labor market experience and dummies for 6 six educational categories: incomplete and complete primary, high school and college education. The marginal returns to a completed college education are calculated compared to a completed high school education, and the marginal returns to a high school education, are calculated compared to completed primary education.

The returns to college fell for women and grew slightly for men during the 1980s and then increased strongly in the 1990s. The returns to a completed college education

[42] We use the terms "return to education" and "educational wage premium" as synonyms in this essay. The coefficients of educational variables estimated in standard Mincer wage regressions are potentially biased due to omitted human capital variables, such as ability, which may be correlated with education. While the causality from education to earnings in undebated, one must thus be cautious about strong inferences about the magnitude of effect, i.e., the actual return. Still, it has become common to refer to the education coefficient in any statistical earnings model as the "(Mincer) returns to education" (Card 1999). An extended discussion on the causality of education on earnings is beyond the scope of this paper. For a more detailed analysis of changes in the returns to education in Argentina, see Giovagnoli, Fiszbein, Patrinos (2005) and Margot (2001).

[43] Marginal returns to education are estimated by specifying standard Mincer wage regressions, where the dependent variable is log hourly wages from principal occupations for adult full-time workers. Regressors include educational dummies and potential labor market experience calculated as age minus years of education minus 6.

compared to a completed high school education were 0.59 for men in 1980, 0.66 in 1992, and 1.03 in 2002. For women the returns of college education were 0.53 in 1980, falling to 0.38 in 1992 and rising to 0.79 in 2002.

The distribution of educational attainment in the labor force has also changed,. Overall, there was a steady educational upgrading in the population, as shown in Table 5-2. Between 1980 and 2002, the fraction of full-time workers with a high school or college degree of each gender doubled, going from 23 to 48 percent for men and from 34 to 66 percent for women.

Table 5-2. Shares of Workers by Educational Groups (in Percent)

MEN	Primary Incomplete	Primary Completed	High School Incomplete	High School Completed	College Incomplete	College Completed	All
1980	19.7	37.7	19.3	11.1	6.8	5.4	*100.0*
1986	15.3	33.3	21.1	13.8	7.9	8.5	*100.0*
1992	9.1	33.4	21.8	17.8	8.6	9.4	*100.0*
1998	6.9	28.5	24.1	18.6	12.2	9.7	*100.0*
2002	6.2	23.3	22.6	19.4	13.4	15.0	*100.0*
WOMEN	**Primary Incomplete**	**Primary Completed**	**High School Incomplete**	**High School Completed**	**College Incomplete**	**College Completed**	**All**
1980	16.0	32.7	17.5	18.7	7.4	7.7	*100.0*
1986	13.8	28.5	18.0	21.6	8.2	9.8	*100.0*
1992	6.6	28.1	17.3	24.9	9.7	13.5	*100.0*
1998	5.4	20.7	17.4	24.1	15.8	16.6	*100.0*
2002	3.7	18.5	12.4	28.1	17.6	19.8	*100.0*

Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly.

The returns to labor market experience have also evolved, but with diverging patterns for men and women. As Table 5-3 shows, the returns to experience fell for men between 1980 and 1992, and then rose back to 1980 levels by 2002. The return to experience evaluated at 20 years was 0.77 in 1980, fell to 0.69 in 1992 and rose back to 0.77 in 2002. In contrast, returns to experience grew constantly for women over the two decades. Evaluated at 20 years, returns increased from 0.53 in 1980 to 0.61 in 1992 and 0.73 in 2002.[44]

[44] It is recognized that "potential labor market experience" measured as *age* minus *years of education* minus *school entry age* most likely overstates actual labor market experience more for women than for men due to child bearing and the traditional division of labor in the family. This results in estimated returns which are biased downwards. See e.g. Blau and Kahn (1997).

Table 5-3. Returns to Labor Market Experience

		Coefficient on		Return to Experience Evaluated at		
	Year	Experience	Squared Experience	10 Years	20 Years	30 Years
MEN	1980	0.039	-0.00055358	0.384	0.769	1.153
	1992	0.035	-0.00049999	0.345	0.690	1.035
	2002	0.039	-0.00046859	0.385	0.771	1.156
WOMEN	1980	0.027	-0.00045122	0.265	0.531	0.796
	1992	0.031	-0.00052025	0.305	0.610	0.914
	2002	0.037	-0.0004366	0.366	0.731	1.097

Note: Returns to experience are estimated in standard Mincer wage regressions, in which the dependent variable is log hourly wages from principal occupations for adult full-time workers with one single job. Regressors include a quadratic in potential labor market experience and dummies for 6 six educational categories: incomplete and complete primary, high school and college education.

Table 5-4 depicts changes in the distribution of labor market experience among full-time workers. While Argentina experiences a gradual ageing of the population, the distribution of labor market experience remains roughly constant between 1980 and 2002. This fact may be due to longer periods of education.

Table 5-4. Shares of Workers by Groups of Years of Experience (in Percent)

MEN	[0 - 10)	[10 - 20)	[20 - 30)	≥ 30	All
1980	18.5	23.1	21.6	36.8	*100.0*
1986	17.4	25.2	22.9	34.6	*100.0*
1992	22.4	23.9	21.1	32.7	*100.0*
1998	21.7	23.4	22.3	32.6	*100.0*
2002	18.9	24.1	24.0	33.0	*100.0*
WOMEN	**[0 - 10)**	**[10 - 20)**	**[20 - 30)**	**≥ 30**	**All**
1980	30.6	23.1	18.9	27.4	*100.0*
1986	27.2	23.6	18.9	30.3	*100.0*
1992	29.9	19.0	23.7	27.5	*100.0*
1998	30.7	21.5	19.5	28.3	*100.0*
2002	28.2	23.7	20.9	27.1	*100.0*

Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

5.2.2 Potential Explanations for Changes in Earnings Inequality

There are several possible explanations for the rise in inequality that has taken place in Argentina in the 1990s, in particular the increase in the returns to education.

Candidate explanations include institutional changes, such as changes in the minimum wage and union membership, trade-related changes in production, and skill-biased technological change (SBTC). We briefly consider whether institutional changes or trade may be responsible and then describe the evidence for SBTC in more detail.

A. Institutional Changes

Some have argued that the increase in wage inequality observed in the 1980s in the United States was driven largely by dramatic declines in unionization rates and the real value of the minimum wage (see, for example, Dinardo, Fortin, and Lemieux 1996). Although effects of unions and the minimum wage on employment could go in either direction, it is generally expected that unions and the minimum wage reduce wage inequality among the employed by boosting the wages of those in the lower part of the wage distribution.

While data on unionization rates in Argentina is not collected consistently, estimates from household surveys show that union membership among non-agricultural salaried workers in Greater Buenos Aires increased only slightly from 45 percent in the beginning of the 1980s to 49 percent in 1990. Over the next decade, it fell to 42 percent in 2001 (Marshall 2005). The relationship between union membership, policies, and labor market outcomes is complex, particularly in a country like Argentina where unions have a strong voice in political decisions. But narrowly examined, the fairly small drop in union membership is not large enough to explain the increases in wage inequality that took place in the 1990s.

Changes in minimum wage also have the potential to influence wage inequality. The real value of the minimum wage greatly eroded during the period of hyperinflation of the late 1980s and early 1990s. Because it dropped to a point below the wages of essentially all workers, the minimum wage cannot have contributed to changes in inequality between the early and late 1980s. However, it is conceivable that the minimum wage during the 1990s did have some effect. The nominal wage was increased from 97 pesos to 200 pesos in 1993 and remained there until 2003. Between 1992 and the end of the fixed exchange rate in December 2001, the minimum wage remained essentially unchanged in real terms.

As a first order effect, the increase in the minimum wage that took place in 1992 would be expected to decrease wage inequality, at least among formal salaried workers. As Maloney and Mendez (2003) point out, the minimum wage can have complex effects on the wage distribution, beyond those on formal salaried workers near the minimum wage. In many countries, the minimum wage has both "numeraire" and "lighthouse" effects that spill over to the informal sector. The numeraire effect is the bunching of wages at round multiples of the minimum wage, due to the fact that the statutory minimum wage is often used as the numeraire for wage negotiations. The lighthouse effect refers to the concentration of informal workers (for whom the minimum wage is not enforced) at the minimum. Using 1998 EPH data, Maloney and Mendez find strong evidence of both effects in Argentina. Likewise, Khamis (2007) examines the effects of changes in the minimum wage 1993 and 2004 on wages and finds positive effects on both formal and informal wages, with a larger effect for informal wages.

Given these effects, it is difficult to determine with certainty what the wage distribution would have looked like with a lower minimum wage. Overall, however, it seems likely that both the direct effect of the minimum wage increase and the numeraire and lighthouse effects tended to raise the wages of those in the lower part of the distribution, reducing inequality even while overall wage inequality increased.

B. Trade

Time trends at first glance suggest that widening inequality may be due to the trade liberalization that took place over the course of the 1990s, and a wide international literature has considered the possible effects of trade opening on wage inequality and the returns to skill. Theory suggests that liberalization towards countries with large numbers of unskilled workers may increase the gap between wages of the skilled and unskilled. Porto (2003) shows evidence that a substantial portion of Argentine imports are substantially unskilled labor-intensive, which lends some credibility to the hypothesis that trade is behind the increase in returns to skill. Using a Computable General Equilibrium approach, Cicowiez (2003) finds that declining import tariffs increased the gap between skilled and unskilled workers only to a negligible amount, explaining between 0 and 6 percent of the change, depending on model and assumptions. A more direct test of the hypothesis is carried out by Galiani and Sanguinetti (2003) by testing whether sectors where import penetration deepened are also the sectors where a higher increase in wage inequality is observed. They find some evidence that this is the case but conclude that trade deepening can only explain a small portion of the observed rise in wage inequality.

C. Technological Change

Skill-biased technological change denotes the phenomenon by which relative wages may change in a country due to the adoption of new technologies. If such technologies are complementary to skills, then workers with these skills will benefit from increased productivity of these skills and consequently increased returns or compensation of these skills. The wage distribution will spread as the workers without the complementary skills are less in demand and their relative wages will fall, resulting in increased wage inequality.[45]

A line of literature for the U.S. starting with Katz and Murphy (1992) looks at SBTC in a supply and demand framework. The approach in these studies is to divide employment into various cells, e.g. by age-gender-education, and examine the relationship between changes in wages and employment by cell over time, applying assumptions about the elasticity of substitution between workers in different groups. The SBTC literature has been criticized on a number of grounds (see for example Card and Dinardo, 2005). The most substantial critique is that the effect of SBTC is always a residual out of a model-based estimation, and the estimates tend to be highly sensitive to the particular assumptions that go into the model. This is because the "facts" to be explained by the analysis are the changes in the cell means. The presence of technological change is inferred by a failure of the model to rationalize the co-movements of wages and employment for different groups over the sample period.

Other studies apply a variance decomposition analysis over time. The objective is to split up the variance over time into its components, the variance within and between groups of the same education and experience. Changes in the returns to observed skills, such as education, change the distance between the mean wages of different population sub-groups. An increase in the returns to higher education will drive the sub-group means further away from each other, thus increasing earnings dispersion. SBTC might be the

[45] For an extended discussion on this, see Acemoglu (2002).

reason for such an increase in returns to education. Since very few individual skills are observed in the data, individuals with heterogeneous unobserved skills will look alike to the econometrician. If the returns to some unobserved skill change, this will be noted as changes in the earnings dispersion within sub-groups, the residual variance. With certain assumptions in the decomposition process one can infer changes in the returns to unobserved skills. People who argue for SBTC have also claimed that SBTC may also change unobserved skill returns, and that changes in these returns may be taken as indications for SBTC.

Lemieux (2006) points out the role that changes in the composition of the workforce have for the residual variance. Taking composition effects appropriately into account, he finds that an increase in returns to unobserved skill may have occurred in the United States in the 1980s but did not in the 1990s when technological progress is widely believed to have taken place. The fact that the returns to unobserved skill increase only in the 1990s is also incongruent with the consistent increase in the returns to education over both the 1980s and 1990s. Overall, Lemieux concludes that the pattern of changes in the returns to unobserved skill in the United States does not lend support to the SBTC hypothesis.

In the literature in the U.S. and other countries, the lack of evidence for other explanations is interpreted to imply that SBTC may be behind increases in the returns to education. The same holds true for Argentina. Several reviews suggest that changes in technology are the proximate cause of changes in returns to education in Argentina. Giovagnoli, Fiszbein, and Patrinos (2005) suggest that increased demand for skills may have driven the increasing returns to education observed in the 1990s. Analysis in World Bank (2003) also shows that the patterns observed for that decade are consistent with skill-biased technical change. Acosta and Gasparini (2007) show that the wage premium for a college education increased more in manufacturing industries with higher rates of physical investment. They also find that this premium grew more in sectors which faced strong import competition.

In an extensive analysis of labor market data from Gran Buenos Aires, Gasparini (2003) presents many pieces of evidence in favor of SBTC as an explanation for the increase in inequality which Argentina experienced in the 1990s. He especially contrasts the economically frustrating experience of import substitution industrialization until the end of the 1980s with the significant productivity increase experienced in the 1990s through reforms and international market integration. Measures of technological progress are hard to obtain, but increases in private investment as a proportion of GDP, a fall in the average age of the capital stock, and a strong increase in the imports of capital goods are indirect evidence of the incorporation of new technologies in the Argentine economy after 1991. Given the parallelism of reforms and the immediate nature of liberalization and opening of the economy to international competition, this might be regarded a "true technological shock" to Argentina.

In this context, trade and technological change may clearly be connected. Trade opening enables the import and adoption of technology-intensive foreign capital and goods. However, when comparing the two direct channels, import penetration of abundant-skill intensive goods and technological change, several studies, including for Argentina, underline the dominance of the technology channel (see Gasparini 2003; Acosta and Gasparini 2004).

5.3 Theoretical Framework for Earnings Inequality and Returns to Skill

This section presents the theoretical fundamentals of the analysis of changes in the distribution of wages, incorporating the role of changes in the distribution and remuneration of skills, such as education and experience. We follow the methodology employed by Lemieux (2006) for the United States. The approach can be considered a generalization of the Oaxaca-Blinder decomposition of means to the case of an entire distribution.

In his seminal work, Mincer (1974) laid the foundation for a vast strand of research on human capital earnings. He specified the earnings function

(1) $$\log w_{it} = \alpha_t + X_{it}\beta_t + \varepsilon_{it}$$

where w_{it} is the hourly wage rate of individual i at time t, α_t is a constant, X_{it} is a vector of observed personal characteristics, β_t is a coefficient vector, and ε_{it} is the standard regression residual. Personal characteristics usually include a person's education, either in years of schooling or in a vector of dummies for educational attainment, and a quadratic of age or alternatively potential labor market experience. X_{it} can be understood as a distribution of human capital and β_t as its price. ε_{it} contains the unexplained portion of the wage, which is usually quite large due to the vast amount of personal characteristics that a researcher cannot observe in the data. In the literature on returns to unobserved skills, the residual is interpreted as the true residual (including measurement error) μ_{it} plus the product of the return p to unobserved skills at time t with the unobserved skill vector e of individual i:

$$(2) \qquad \varepsilon_{it} = p_t e_{it} + \mu_{it}$$

The variance, as a standard measure of dispersion, of wages is thus

$$(3) \qquad V_t = \beta_t^{'} \Omega_{x,t} \beta_t + \sigma_t^{2} \ ,$$

where $\Omega_{x,t}$ is the variance-covariance matrix of X_{it}, and σ_t^2 is the variance of the error term. Changes in the variance over time can thus be caused by several factors: (a) changes in the distribution of observed characteristics X_{it}, (b) changes in the returns to observed skills, (c) changes in the distribution of unobserved characteristics e_{it}, (d) changes in the returns to unobserved skills, or (e) changes in measurement error.

For equation (2) to have some empirical content, it is necessary to impose some assumption on the distribution of skills. Since both unobserved skill and the returns to unobserved skill are "unobserved," *some* assumption is needed. The usual assumption is that the distribution of unobserved skills among workers with the same observed skills is stable over time.[46] In other words, the *conditional* distribution function does not vary over time:

[46] This assumption is used in Juhn, Murphy and Pierce (1993), Chay and Lee (2000), and Lemieux (2006).

(3) $F_t(e_{it} \mid X_{it}) = F_t(e_{it} \mid X_{it})$ for all time periods *t*.[47]

Note that the stronger assumption sometimes implicitly used in the literature, which is that the *unconditional* distribution of unobserved skills is stable over time, is clearly incorrect. It is well established in both the theoretical and empirical literature that heteroskedasticity is pervasive in wage regressions, and wage dispersion increases with both education and experience. Consequently, changes in the composition of the workforce, i.e. in the relative size of education-experience groups, will change the unconditional distribution of unobserved skills, even with no change in the return to unobserved skills.[48]

Although the issue is sometimes ignored, it is crucial to control for composition effects when considering the changes over time in the returns to unobserved skill. The role of composition effects is illustrated by considering the variance of wages. Consider the case where observed skills, X_{it}, are divided up into *j* cells. Then, the unconditional variance of unobserved skills is the weighted sum of the conditional variances for the *j* subgroups. The weights are simply the shares, θ_{jt} , of workers in experience-education group *j* at time *t*:

(4) $Var(e_{it}) = \sum_j \theta_{jt} Var(e_{it} \mid j)$.

Give the assumption that the conditional variances are stable over time, this equation can be written as follows:

(5) $Var(e_{it}) = \sum_j \theta_{jt} \sigma_j^2$,

where $Var(e_{it} \mid j) = \sigma_j^2$ for all *t*.

[47] As pointed out by Lemieux (2006), this assumption may be problematic e.g. if there are cohort effects: younger cohorts could have a different distribution of unobserved skills conditional on education, e.g. due to change in school quality or educational content.
[48] This point was raised by Lemieux (2004) and is also explained by Card and Dinardo (2005).

Note that because the conditional variances, σ_j^2, are different for every skill group, changes over time in the shares in each group (e.g. increased education levels or aging of the workforce) will also change the unconditional variance of unobserved skills.[49]

The residual variance of wages, which is what can be estimated in wage regressions, is given by taking variances of equation (2) – ignoring measurement error – and substituting in equation (5):

(6) $$Var(\varepsilon_{it}) = Var(p_t e_{it}) = Var(p_t) * Var(e_{it}) = p_t^2 \sum_j \theta_{jt} \sigma_j^2 .$$

What we are interested in is how the price of unobserved skills, p_t, may have changed over time. A change in the residual variance of wages can only be interpreted as a change in the price of unobserved skills if the skill shares in the workforce, θ_{jt}, are held constant over time. Note again that the actual skill shares tend to change over time, as education levels increase and the workforce ages.

Some empirical papers ignore this problem, and treat changes in the residual variance of wages as being equivalent to changes in the price of unobserved skills. There are, however, multiple ways to correct for the problem. One way is to calculate the residual variance at counterfactual values of the shares θ_j^* that are held constant over time. We can rearrange (6) as follows:

(7) $$Var(\varepsilon_{it}) = \sum_j \theta_{jt} (p_t^2 \sigma_j^2)$$

If we hold the shares constant, the variance becomes the following:

(8) $$Var(\varepsilon_{it})^* = \sum_j \theta_j^* (p_t^2 \sigma_j^2)$$

The within-group variances, $p_t^2 \sigma_j^2$, can be computed for each skill group *j,* if the number of skills groups is small enough relative to the sample size that there are substantial numbers of observations in each skill group. The overall variance at the

[49] This is illustrated in Card and Dinardo (2005) for the simplest case, with just two skill groups.

counterfactual shares can then be calculated, using shares either in the initial year, the final year, or the average of the two. The variance can be calculated using all three methods as a sensitivity test. Changes in this "counterfactual" variance provide an estimate of changes in the returns to unobserved skill.

A more convenient way to correct for composition changes is to re-weight the data for the purposes of calculating the residual variance so that the distribution and prices of observable skills at time *t+1* is identical to the distribution and price of skills at time *t*. The re-weighting procedure is in the spirit of Dinardo, Fortin, and Lemieux (1996) and is described in Lemieux (2002) and Lemieux (2004). The advantages to the re-weighting procedure are two-fold. First, it can be applied even when the data is divided into fine experience-education cells. Second, it provides a whole counterfactual wage distribution and thus makes it possible to compute measures of residual wage dispersion other than the variance, e.g. the ratios between different percentiles of the residual distribution.

It should further be noted that measurement error is an additional factor which may, if its extent changes over time, introduce a change in residual variance which is unrelated to unobserved skills or returns. We already mentioned the case of hyperinflation, where measurement error most likely renders any analysis useless. Our solution to this problem is to consider years for comparison which are less affected by inflation. This is most relevant for the 1980s, where we consider 1980 and 1986 the most appropriate base years. Apart from that we have no means of analyzing if and how measurement error has changed over time in the EPH and thus assume it constant.

5.4 Data and Estimation Issues

The data used for this analysis is the household survey Encuesta Permanente de Hogares (EPH) of Argentina which has been carried out by Argentina's statistical office (INDEC) since 1972 and is used as the primary source of generating official unemployment rates. The survey includes comparable labor market information from 1980 through 2003 for the province of Gran Buenos Aires (GBA). The GBA sample encompasses the capital city of Buenos Aires and the surrounding Province of Buenos

Aires. According to the Argentine Census, 46 percent of the Argentinean population lived in this area. As Argentina is mostly urban, trends observed in Buenos Aires are often considered representative for Argentina as a whole.

More urban centers of Argentina were later added to the sample over time, totaling 28 major provincial cities in the most recent incarnations of the survey. There is data with comparable coverage since 1992 for 16 main urban conglomerates in Argentina (henceforth ARG16). Until 2003, the survey was conducted on a semi-annual basis (May and October) before the questionnaire and methodology changed substantially.

We investigate the time series for GBA from 1980-2002, always using the October round of the survey[50]. For the wage analysis we focus on real hourly wages of workers with one single job only as reported in the EPH questionnaire[51]. To convert nominal wages into real wages we use INDEC's historic general consumer price index (IPC) for Gran Buenos Aires and deflate all values to constant October 2000 Pesos.

To underline the explanatory power of the results from the smaller GBA sample, the decomposition analysis is also carried out using the ARG16 sample from 1992-2002 as a robustness check. For the analysis, the sample of urban centers is not continuously expanded to 28 cities as survey coverage increases over time. This is because changes in the survey's coverage can have substantial effects on the residual variance induced by geographical differences, which we cannot observe. This may be the case even if there are no important changes in the observable means. Regional variation in the ARG16 sample is accounted for by adjusting all incomes to the level of GBA, using a one-time comparison of price levels in 2001. This method effectively incorporates the assumption that relative regional price differences have not changed over time. However, due to the convertibility regime from 1991 to 2001 and the according price stability this assumption may be justified for most years of the ARG16 sample, yet arguable for later years.

---

[50] The May round of 2003 could be used to expand the data by another half a year, however in an analysis of variance this might be rather misleading due to seasonality effects on employment and wages. Data from INDEC clearly shows that there is considerably higher economic activity in May than in October.

[51] To avoid effects stemming from changes in the incidence of multiple-job holders this paper focuses on wages from the principal occupation, only. In order to do that one has to discard workers with more than one occupation in order to establish consistency of the data series over time. Before 1995, hourly wage data is only available for those workers with one single job. Even though this may be a minor point, to our knowledge this adjustment to guarantee consistency has not been done yet in any empirical research using EPH data.

Data inspection reveals a strong spike in all wage dispersion figures centered around 1989, the worst year of hyperinflation in Argentina (see Figure 5-6 and Figure 5-7, appendix). Prices soared up to nearly 4000 percent annually, which led to the introduction of the Argentine currency board in April 1991. Measurement error is likely to be higher in times of high inflation, if people have to recall their earnings in an environment of constantly changing prices and wages. Second, during hyperinflation, prices and wages change monthly, weekly or even daily. Since surveys cannot be carried out at the same point in time for the whole sample, sequenced interviewing will introduce an upward bias to the wage variance in times of high price volatility and wage contract turnover

Thus, the figures for the 1980s must be analyzed with caution. Using a base year with a bloated wage variance might lead to wrong conclusions of variance changes over time. What matters for the data quality from periods of high inflation is not only the yearly inflation but also the inflation figures from the month of interviewing. We use 1980 as base years, as there was moderate inflation during both the whole year and in the survey month of October.

We apply the reweighing methodology to analyze changes in the residual variance over time against a base year by re-weighting the observations of the more recent year. The educational and demographic distribution of the Argentine labor force has changed noticeably since the 1980s. In particular, the overall improvement in educational attainment may have increased wage dispersion over time.

The decomposition is carried out stepwise, following Lemieux (2002): first, a counterfactual wage distribution is generated, using the later year's observable skill distribution and the base year's estimated coefficients on observed skills. The difference between the inequality indicators of the final year and those of the counterfactual distribution can be attributed to changes in the returns to observed skills. In a second step, the counterfactual distribution is re-weighted as detailed above. The difference between inequality indicators of the two distributions is ascribed to changes in the skill distribution in the population. Finally, the difference between the distributional indicators of the base year and the counterfactual distribution using both, base year weights and returns, is the effect of changing returns to unobserved skills.[52]

5.5 Analysis of Earnings Inequality in Gran Buenos Aires

Table 5-5 and Table 5-6 present variance decomposition results for men and women, for the periods 1980-1992 and 1992-2002 which we focus on separately since they correspond to two different economic policy regimes.[53] Four measures of wage dispersion are depicted: the variance of log hourly wages, and the 90:10, 50:10 and 90:50 percentile ratios. The first three rows of each table show values in the base year and the final year, and the absolute changes. Rows four to six split up the change into three components and show how much of the overall change was caused by changes in each of the components: the returns to observed skills, the composition of observed skills in the workforce, and the returns to unobserved skills.

5.5.1 Period 1980-1992

Between 1980 and 1992, wage dispersion decreased for both men and women (see Table 5-5). For men, the variance of log wages dropped from 0.48 to 0.45 and the 90:10 ratio from 5.5 to 5.0. Improvements in the education level of the workforce tended to increase the overall variance of wages, due to the fact that groups of workers with higher education have higher within-group wage dispersion. On the other hand, decreases in the returns to observed skills tended to decrease in the overall variance. The overall effect of falling returns to high school education and falling returns to experience dominates the inequality-increasing effect of rising returns to college education, which only improves the wages of the college-educated minority of the population. The results show that at the same time inequality among men was lowered due to decreasing returns to unobserved skill.

[52] Decomposition results switching the order of the first two steps are qualitatively similar to the base case and can be obtained from the authors upon request.
[53] Decomposition results for the whole period 1980-2002 are presented in Table 5-14 of the appendix.

Table 5-5. Decomposition Results, 1980-1992

Panel A. Men

	Var logwage	90/10	50/10	90/50
1992	0.449	5.036	1.892	2.662
1980	0.481	5.488	2.091	2.625
Change	-0.032	-0.452	-0.199	0.037
Components of change:				
Δ Returns to Observed Skills	-0.040	-0.496	-0.084	-0.138
Δ Composition of Observed Skills	0.063	0.724	0.114	0.217
Δ Returns to Unobserved Skills	-0.055	-0.681	-0.229	-0.043

Panel B. Women

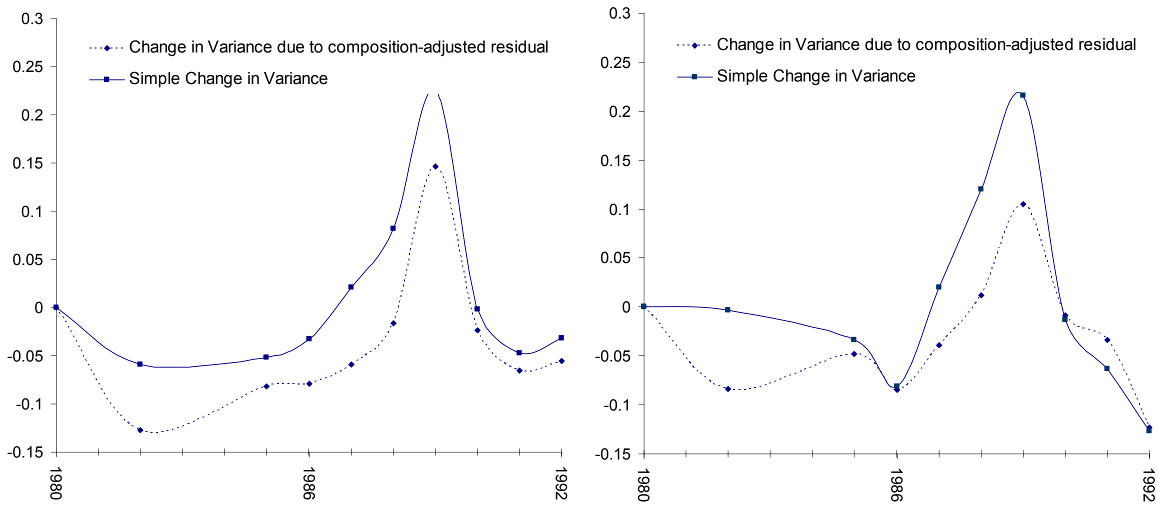
	Var logwage	90/10	50/10	90/50
1992	0.338	4.102	1.820	2.253
1980	0.465	5.065	2.101	2.411
Change	-0.127	-0.963	-0.281	-0.157
Components of change:				
Δ Returns to Observed Skills	-0.094	-1.150	-0.180	-0.372
Δ Composition of Observed Skills	0.090	1.115	0.184	0.348
Δ Returns to Unobserved Skills	-0.123	-0.929	-0.285	-0.133

Note: Estimations for single-job workers age 15 or above working 30 or more hours weekly

Figure 5-2 depicts the change in overall variance for men and women, and to what extent the change is a result of changes in the composition-adjusted residual variance.

Figure 5-2. Change in Variance and Residual Variance, 1980-1992, Men and Women

Panel A. Men **Panel B. Women**



Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

In 1980, the male wage distribution showed a higher 50:10 than 90:50 ratio. This imbalance is increased in the 1980s as the 50:10 ratio falls (from 2.1 to 1.9) while the 90:50 ratio rises slightly (2.6 to 2.7). Composition effects, i.e., the increase in educated workers, increased inequality in the upper half. Changing returns to observed skills decrease inequality more in the upper half, mostly due to falling returns to experience. In contrast, the changes in the residual distribution decrease the 50:10 ratio more strongly, i.e., returns to unobserved skills affect the lower part of the distribution more.

For women, the variance fell from 0.47 to 0.34 and the 90:10 ratio from 5.1 to 4.1. The drivers of the decrease are the same as for men: composition effects contributed to an increase in wage variance which was counteracted by falling returns to observed skill, where lowered returns to higher education dominate the effect of rising returns to experience. At the same time, decreasing returns to unobserved skill also lowered inequality among women.

As for men, the changes of the 1980s increased the imbalance of the female distribution which has a higher inequality in the upper half: the 90:50 ratio fell less (2.4 to 2.3) than the 50:10 ratio (2.1 to 1.8). The reason is the stronger decrease in returns to unobserved skill in the lower half of the distribution. Composition effects increased the 90:50 more but changing returns to observed skills counteracted this effect: women in the labor market are on average more educated than male workers (see Table 5-2) but have less experience (Table 5-4), so they are overall more hurt by falling returns to education than helped by increasing returns to experience.

To sum up, the patterns of the 1980s are similar for men and women: between 1980 and 1992, a combination of falling or stagnant returns to higher education and changing returns to experience caused the between-group variance to decrease as the means of different education-experience groups moved closer together. Composition effects via educational upgrading of the workforce increased inequality as higher educational groups tend to have higher within-group wage variances. The results also suggest that the returns to unobserved skill decreased in the 1980s in Argentina, in line with returns to observed skills.

5.5.2 Period 1992-2002

The 1990s show a substantially different picture from the 1980s (see Table 5-6). Male wage variance increased strongly, from 0.45 to 0.68, and the 90:10 ratio jumped from 5.0 to 7.9. The strongest drivers of this increase were changes in the returns to observed skills. To a smaller degree, composition effects and changing returns to unobserved skill also contributed to the increase. Figure 5-3 shows to what extent changes in the variance can be explained by changes in the composition-adjusted residual.

Table 5-6. Decomposition Results, 1992-2002

Panel A. Men

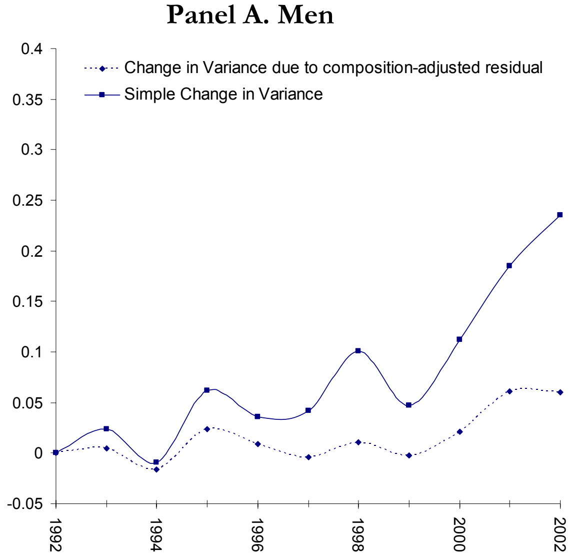
	Var logwage	90/10	50/10	90/50
2002	0.684	7.882	2.361	3.338
1992	0.449	5.036	1.892	2.662
Change	0.235	2.846	0.469	0.676
Components of change:				
Δ Returns to Observed Skills	0.139	1.919	0.059	0.748
Δ Composition of Observed Skills	0.036	0.557	0.050	0.190
Δ Returns to Unobserved Skills	0.060	0.370	0.361	-0.262

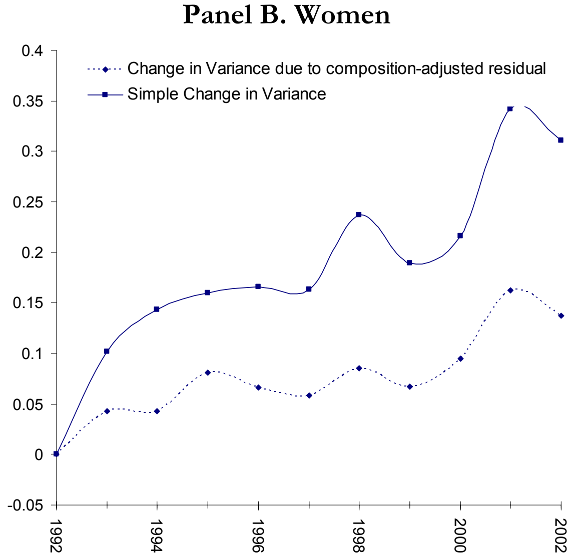
Panel B. Women

	Var logwage	90/10	50/10	90/50
2002	0.649	7.361	2.577	2.856
1992	0.338	4.102	1.820	2.253
Change	0.311	3.259	0.757	0.603
Components of change:				
Δ Returns to Observed Skills	0.175	2.328	-0.070	0.955
Δ Composition of Observed Skills	-0.002	-0.229	-0.156	0.024
Δ Returns to Unobserved Skills	0.137	1.161	0.983	-0.376

Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

Figure 5-3. Change in Variance and Residual Variance, 1992-2002, Men and Women





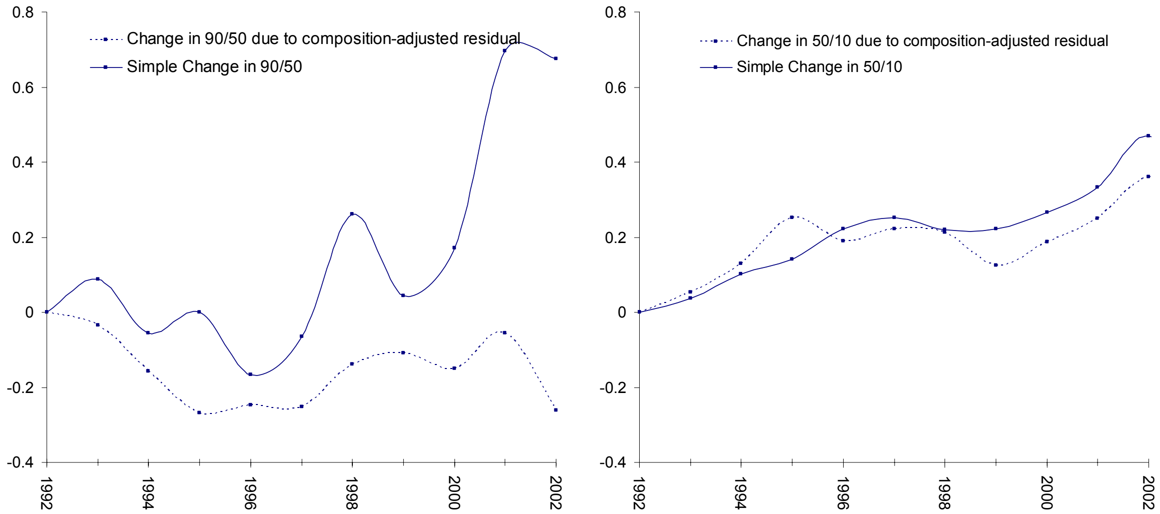
Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

For men, the 1990s also saw widening earnings inequality, measured both in terms of the 90:50 and 50:10 ratios. For changes to the 90:50 ratio, most of the inequality increase was due to changing returns to observed skill. The returns to university education grew much more than returns to secondary education, and the mean wage of the highly-educated is strongly shifted upwards. Also, returns to experience increased, favoring more experienced workers who already have a higher within-group variance. Composition effects played a much smaller role, and returns to unobserved skills decreased in the upper distribution half. On the contrary, in the lower half of the distribution, the changing returns to education explain only about 13 percent of the inequality increase. However, increasing returns to unobserved skills explain almost 80 percent of the increase in the 50:10 ratio. The contribution of the returns to unobserved skills to explain changes in the upper and lower half of the distribution are depicted in Figure 5-4.

For women, the variance of wages also increased between 1992 and 2002: from 0.34 to 0.65. The 90:10 ratio grew from 4.1 to 7.4. Composition effects played an almost negligible role for women, and were even slightly negative for the 90:10 ratio. Most of the variance growth was driven by increases in the returns to observed skills, followed by increases in the returns to unobserved skills.

Figure 5-4. Change in 90:50 and 50:10-ratio, 1992-2002, Men

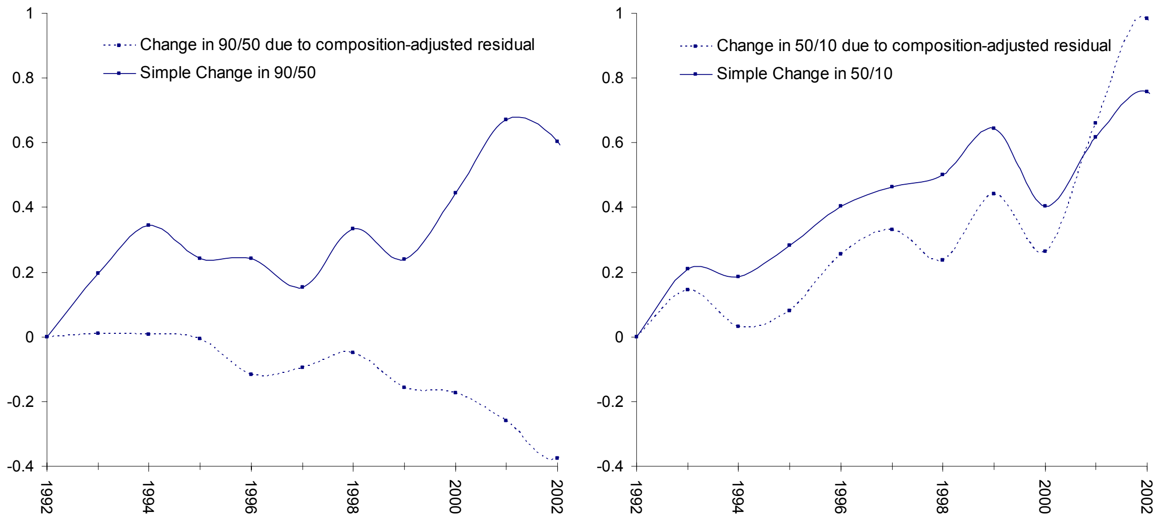
Panel A. 90:50 Ratio **Panel B. 50:10 Ratio**



Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

Figure 5-5. Change in 90:50 and 50:10-ratio, 1992-2002, Women

Panel A. 90:50 Ratio **Panel B. 50:10 Ratio**



Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

The trends of the female wage distribution in the 1990s is very similar to that of men: inequality grew in both parts of the distribution, with the 50:10 ratio increasing from 1.8 to 2.6 while the 90:50 ratio grew from just 2.3 to 2.9. Increasing returns to higher education and experience played the main role in spreading the upper half of the distribution but no role in the lower half. Growing returns to unobserved skills strongly raised the lower half of the distribution but had a negative effect on the upper half, as in the case of men. The contribution of the returns to unobserved skill to explain changes in the upper and lower half of the distribution are shown in Figure 5-5. As for the case of men, the graphical series underlines the considerable difference in both parts of the distribution.

The fact that composition effects of observed skills played only a small or partly even negative role is due to the fact that the within-group variances of experience-education groups (especially educational groups) converged considerably in the 1990s, mitigating the effect of skill-upgrading on the variance. As Table 5-7 shows, the variance between 1992 and 2002 increased within each experience group for men and women. Since the size of experience groups does not change drastically (Table 5-4), the composition effect with respect to these groups is small. On the other hand, the sizes of educational groups change much more over the course of the 1990s (Table 5-2). Here, the within-educational group variance increases strongly for workers with complete or incomplete primary education (Table 5-8). This effect is especially strong within the female workforce. As these groups are shrinking in the population, the composition effects work towards decreasing the variance.

Table 5-7. Variance of Log Wages by Experience Group

MEN	[0 - 10)	[10 - 20)	[20 - 30)	≥ 30	Overall
1980	0.368	0.458	0.509	0.523	0.481
1992	0.345	0.518	0.480	0.424	0.449
2002	0.637	0.643	0.799	0.649	0.684

WOMEN	[0 - 10)	[10 - 20)	[20 - 30)	≥ 30	Overall
1980	0.380	0.519	0.509	0.461	0.465
1992	0.267	0.323	0.451	0.317	0.338
2002	0.389	0.682	0.799	0.719	0.649

Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

Table 5-8. Variance of Log Wages by Education Group

MEN	Primary Incomplete	Primary Complete	High School Incomplete	High School Complete	College Incomplete	College Complete	Overall
1980	0.350	0.342	0.400	0.448	0.369	0.598	0.481
1992	0.284	0.272	0.331	0.396	0.413	0.550	0.449
2002	0.688	0.367	0.299	0.541	0.550	0.422	0.684
WOMEN	**Primary Incomplete**	**Primary Complete**	**High School Incomplete**	**High School Complete**	**College Incomplete**	**College Complete**	**Overall**
1980	0.338	0.311	0.404	0.289	0.336	0.596	0.465
1992	0.123	0.248	0.217	0.302	0.280	0.471	0.338
2002	0.275	0.617	0.457	0.364	0.279	0.460	0.649

Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

5.5.3 Summary

The decomposition shows that 1980s and 1990s, returns to education and to unobserved skills move in tandem. It thus seems plausible that both phenomena might be driven by the same underlying processes in the case of Argentina.

In line with the analysis, the changes observed for Argentina in the 1980s may have occurred via the technology channel. It is not likely that a form of 'negative technological change' was at work, the explanation may rather lie in the level of technology present, and the supply of and demand for skills: the 1980s were characterized by constant waves of crises and instability. Average GDP growth 1980-1990 was around minus one percent. Over the same period, capital formation was reduced by 50 percent in real terms.[54] There was certainly no positive technology shock during this period – most likely, the technological level was stable, if not decreasing due to lack of replacement and maintenance in times of crises. If, in line with educational upgrading, "technological skills" were improved in the population[55] and the stock of technology to operate deteriorated, we would expect increasing supply of, and falling demand for those skills and consequently falling returns.

The SBTC hypothesis predicts that technology shocks will spread the earnings distribution by increasing the returns to skills that are complementary with the

[54] Own calculations based on World Development Indicators.
[55] Remember that we assume a stable distribution of unobserved skills, conditional on observed skills, in order to interpret changes in the residual as changes in the returns to unobserved skills.

technology. These skills should partly be reflected in educational attainment.[56] The relevant skills which are uncorrelated with education will also experience an increase in demand and in their returns. In the 1990s, the returns to both observed skills, especially university education, and unobserved skills worked strongly towards increasing the wage variance. As noted earlier in this paper, the literature on the United States shows that the returns to observed and unobserved skill have not grown at the same time, casting doubt on the hypothesis that SBTC is behind both changes. In Argentina, however, the two have evolved together. This is compatible with the hypothesis that SBTC is behind both sets of changes.

In the male and female wage distribution, increases in the variance occurred in both the upper and lower halves of the earnings distribution. The drivers, however, are fundamentally different. The increase in the 90:50 ratio is mostly caused by increasing returns to observed skills, especially college education, as college-educated individuals are concentrated in the highest deciles of the wage distribution. On the other hand, the returns to high school education did not increase strongly. As mostly workers with lower to middle education populate the lower half of the wage distribution, the effect of returns to education is weak in this part. Instead, the growth of the 50:10 can almost fully be explained by changes in the residual variance. This means that increasing returns to unobserved skills are the driver of the growth in dispersion between the median and the 1st decile. In other words, workers at the bottom and in the middle of the wage distribution must differ with respect to their unobserved skills in a way in which the middle does not differ from the top of the distribution.[57]

5.6 Analysis of the Informal-Formal Wage Gap

This section considers what relevance the findings for the return to unobserved skill may have for the earnings gap between formal and informal salaried workers.[58] One

[56] This also means that they may be partially correlated with educational attainment. This will introduce a bias to the estimated coefficients to education which will partly reflect returns to unobserved skills.

[57] The same patterns with very similar magnitudes of change hold for the much bigger sample of 16 Argentinean urban conglomerates between 1992 and 2002. The results are shown in Table 5-15 of the appendix.

[58] The paper uses a three-way classification of employment: formal employees, informal employees, and independent workers. Independents are defined as the self-employed and those who are owners of micro-enterprises with 5 or fewer employees. Formality is defined in terms of worker benefits, specifically having the

of the puzzles in Argentina is the long-term transformation of its workforce. In Gran Buenos Aires between 1980 and 2002, the fraction of all workers who are informal salaried workers more than doubled, from 15 to 32 percent, while the share of self-employed workers remained roughly constant at around 26 percent. The shift from formal to informal employment happened steadily over the whole period. Table 5-9 documents how this increase divides up between full-time and part-time workers. Even though most of the increase happens among part-time workers, it is also substantial among full-time workers, which are the object of analysis in this paper.

Table 5-9. Changes in the Employment Structure

	Full-time Workers			Part-time Workers		
MEN	Self-Employed	Informal Salaried	Formal Salaried	Self-Employed	Informal Salaried	Formal Salaried
1980	27.3	11.5	61.2	36.4	19.4	44.2
1986	27.0	13.1	59.9	41.9	19.4	38.6
1992	28.4	18.6	53.0	35.1	32.4	32.5
1998	23.0	24.7	52.4	41.0	37.2	21.8
2002	26.8	23.3	49.8	38.6	42.8	18.5
WOMEN						
1980	20.0	16.8	63.2	35.8	27.5	36.7
1986	20.7	19.6	59.7	29.0	35.5	35.5
1992	21.9	24.2	53.9	31.0	33.0	36.0
1998	21.4	25.0	53.6	24.2	43.3	32.5
2002	19.1	24.1	56.7	20.8	50.7	28.6

Note: Population considered are single-job workers age 15 and above. Full-time workers are those working 30 and more hours weekly.

World Bank (2008) documents the evolution of informal employment in Argentina, exploring possible explanations for the steady increase in rates of informal employment over time. Possible causes include macroeconomic policy and privatization, economic structure and demographic change, trade and technological change, labor regulations and institutions, as well as tax evasion, enforcement and weak public confidence. In a simply supply and demand framework, decreasing wages in light of increasing demand for informal workers could be a natural result of large increases in the supply of informal workers. The most likely candidate explanation for such changes in

right to receive a pension, which has been shown to be highly correlated with registration in the social security system (World Bank 2008).

supply would be changes to the labor market induced by structural or demographic changes, such as an augmented entering of women into the labor force. However, in separate analyses, Gasparini (2002) and World Bank (2008) find that structural and demographic changes cannot explain the increase in levels of informal employment. World Bank (2008) also shows that changes in the minimum wage and unionization are very unlikely candidates to explain the decreasing informal-formal wage gap. Additionally, changes in trade patterns since 1980 can explain only a small portion of increases in the size of informal salaried employment (Goni and Maloney, 2007).

Table 5-10. Effect of Returns to Education on Relative Wage Ratios

		Mean Earnings Ratio in 1992	Mean Earnings Ratio in 2001	Simulated Mean Earnings Ratio in 2001 with 1992-coefficients	Percent Change Explained By Change in Coefficients
Informal / Formal	Men	0.77	0.56	0.58	10.1%
	Women	0.77	0.51	0.62	40.7%
Self-employed / Formal	Men	1.15	0.86	0.85	-2.1%
	Women	0.98	0.49	0.59	20.6%

Note: Column four denotes to which degree the change in the relative wage gap can be explained by changes in the returns to education. In order to do that, we simulate a counterfactual wage distribution for the year 2001 by replacing the true returns to education (estimated coefficients on education dummies) that year with the returns in 1992. The Mincer equations that are used to estimate the returns and generate the simulated wage distributions include only education dummies and a quadratic in experience.

Several hypotheses imply that high levels of informal employment are fundamentally driven by increased demand for informal work arrangements. Such an explanation would be paired most naturally with increased relative wages in the informal compared to the formal sector. However the wages of informal and self-employed workers relative to formal workers have not increased consistently. Between 1980 and 1992, relative wages of informals were indeed rising. However, Table 5-10 shows that in the 1990s, the relative wages of informal workers fell substantially, by approximately 21 percentage points for men and 26 percentage points for women. It is puzzling why relative informal wages fell at the same time that informal employment was stable or even expanding at high levels.

One possibility is that the growing gap between formal and informal salaried wages was driven by increased returns to education, given that formal workers on average are more educated. Table 5-11 and Table 5-12 show the educational structure of the employed population in Gran Buenos Aires and its evolution between 1980 and 2002. Formal workers have higher levels of completed secondary or higher education, compared to informal workers: e.g., in 1992 the comparison is 41 compared to 20 percent for men, and 61 to 29 percent for women.

Table 5-11. Educational Composition of the Workforce, Men

Self-Employed	Primary Incomplete	Primary Complete	Secondary Incomplete	Secondary Complete	Tertiary Incomplete	Tertiary Complete
1980	19.5%	36.9%	20.4%	8.7%	6.9%	7.5%
1986	13.2%	32.2%	23.3%	15.2%	5.9%	10.2%
1992	7.5%	36.7%	21.3%	17.6%	8.7%	8.3%
1998	8.5%	24.3%	23.7%	20.8%	9.4%	13.4%
2002	8.4%	25.6%	21.0%	17.8%	8.6%	18.7%
Informal Salaried	Primary Incomplete	Primary Complete	Secondary Incomplete	Secondary Complete	Tertiary Incomplete	Tertiary Complete
1980	26.4%	41.1%	20.0%	8.1%	2.5%	1.9%
1986	22.4%	37.6%	24.5%	8.8%	4.5%	2.1%
1992	13.3%	38.5%	28.5%	11.2%	6.0%	2.5%
1998	10.0%	35.6%	28.5%	15.1%	7.2%	3.7%
2002	8.2%	28.6%	29.4%	18.2%	7.3%	8.4%
Formal Salaried	Primary Incomplete	Primary Complete	Secondary Incomplete	Secondary Complete	Tertiary Incomplete	Tertiary Complete
1980	19.2%	37.8%	18.9%	12.3%	7.2%	4.7%
1986	15.2%	33.5%	19.6%	13.8%	9.2%	8.7%
1992	8.8%	30.3%	20.5%	19.4%	9.2%	11.9%
1998	5.2%	27.6%	22.9%	19.1%	15.3%	10.0%
2002	4.5%	20.8%	21.4%	20.8%	18.2%	14.2%
Employed Population	Primary Incomplete	Primary Complete	Secondary Incomplete	Secondary Complete	Tertiary Incomplete	Tertiary Complete
1980	20.1%	37.9%	19.5%	10.8%	6.6%	5.2%
1986	15.6%	33.7%	21.2%	13.5%	7.7%	8.3%
1992	9.2%	33.6%	22.2%	17.4%	8.5%	9.1%
1998	7.1%	28.8%	24.4%	18.5%	11.9%	9.2%
2002	6.4%	23.9%	23.1%	19.4%	13.1%	14.1%

Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

Table 5-12. Educational Composition of the Workforce, Women

Self-Employed	Primary Incomplete	Primary Complete	Secondary Incomplete	Secondary Complete	Tertiary Incomplete	Tertiary Complete
1980	27.1%	36.3%	9.8%	14.6%	7.7%	4.4%
1986	23.5%	33.0%	13.8%	17.1%	3.7%	9.0%
1992	10.1%	35.9%	16.0%	23.1%	4.4%	10.5%
1998	11.0%	29.6%	21.2%	16.5%	7.2%	14.5%
2002	6.5%	37.2%	10.1%	25.1%	10.1%	11.1%

Informal Salaried	Primary Incomplete	Primary Complete	Secondary Incomplete	Secondary Complete	Tertiary Incomplete	Tertiary Complete
1980	27.0%	44.3%	14.1%	7.9%	2.8%	3.8%
1986	21.1%	40.1%	22.7%	8.0%	5.6%	2.6%
1992	10.7%	36.7%	24.0%	17.9%	7.0%	3.7%
1998	8.5%	29.3%	19.1%	23.5%	11.4%	8.3%
2002	7.9%	22.8%	22.7%	18.3%	22.6%	5.6%

Formal Salaried	Primary Incomplete	Primary Complete	Secondary Incomplete	Secondary Complete	Tertiary Incomplete	Tertiary Complete
1980	9.8%	28.4%	20.9%	22.7%	8.4%	9.7%
1986	8.2%	23.4%	18.0%	27.7%	10.5%	12.2%
1992	3.5%	21.1%	14.9%	28.7%	12.7%	19.2%
1998	1.9%	13.5%	15.3%	27.6%	20.9%	20.9%
2002	1.0%	10.5%	8.5%	33.5%	18.1%	28.5%

Employed Population	Primary Incomplete	Primary Complete	Secondary Incomplete	Secondary Complete	Tertiary Incomplete	Tertiary Complete
1980	16.2%	32.6%	17.5%	18.6%	7.3%	7.7%
1986	13.9%	28.7%	18.1%	21.6%	8.1%	9.6%
1992	6.7%	28.1%	17.3%	24.9%	9.5%	13.5%
1998	5.5%	20.9%	17.5%	24.2%	15.6%	16.4%
2002	3.7%	18.5%	12.2%	28.2%	17.7%	19.7%

Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

As the group of formal and informal workers differs in their educational composition, and returns to education shifted over time, the combination of these phenomena might largely explain the changes in relative wages between the groups. We examine the degree to which the change in the relative wage gap can be explained by changes in the returns to education. In order to do that, we simulate a counterfactual wage distributions for 2001 by replacing the true returns to education (estimated coefficients on education dummies) that year with the returns in 1992. The Mincer equations that are used to estimate the returns and generate the simulated wage distributions include only education dummies and a quadratic in experience. We use the year 2001 as the final year in order to abstract from the short-term drop in returns to education that occurred during the economic crisis in 2002. The counterfactual wage distribution for 2001 using returns from 1992 shows that changes in the returns to education only explain 10 percent of the drop in relative wages for men (Table 5-10). For women, changes in the returns play a bigger role and explain 41.5 percent of the drop in relative informal wages.

The decline in the relative wages of informal workers may be linked to the decline in the returns to unobserved skill. The analysis presented in the previous section shows that changes in the returns to unobserved skill affected chiefly earnings inequality below median wages, i.e. the 50:10 ratio. Table 5-13 shows that in both 1992 and 2002, informal workers were concentrated at the bottom of the wage distribution. Consequently, changes in the returns to unobserved skill are likely to have had a substantial effect on the gap between informal and formal wages.

Table 5-13. Deciles of the Wage Distribution By Occupational Category, 1992 and 2002

		Share of Workers in Decile										
	MEN	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	All
1992	Self-Employed	7.9	6.2	4.4	7.1	8.8	9.1	9.6	14.3	13.4	19.3	*100.0*
	Informal Salaried	15.4	12.6	14.2	11.1	9.9	10.0	10.3	5.0	5.6	5.9	*100.0*
	Formal Salaried	9.1	10.7	11.0	10.9	10.6	10.4	10.1	9.9	10.1	7.3	*100.0*
2002	Self-Employed	16.3	7.6	9.0	8.9	6.5	13.3	7.3	7.7	9.1	14.2	*100.0*
	Informal Salaried	19.2	20.3	16.3	11.3	7.8	5.7	6.0	3.9	4.2	5.3	*100.0*
	Formal Salaried	2.4	6.2	7.4	10.0	12.8	10.4	13.3	14.1	13.2	10.1	*100.0*

	WOMEN	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	All
1992	Self-Employed	15.9	6.2	8.8	7.6	5.1	9.5	10.2	8.9	5.3	22.3	*100.0*
	Informal Salaried	13.4	10.6	10.3	14.9	11.2	11.2	8.8	6.5	5.7	7.4	*100.0*
	Formal Salaried	6.8	11.0	10.2	8.8	11.1	9.7	10.4	11.7	13.2	7.1	*100.0*
2002	Self-Employed	31.2	15.0	6.3	9.2	4.7	3.7	4.7	8.3	3.9	13.0	*100.0*
	Informal Salaried	13.5	18.3	9.4	15.4	9.9	12.1	12.0	3.1	3.5	2.7	*100.0*
	Formal Salaried	1.8	4.8	11.4	7.9	11.7	11.1	10.8	13.5	14.7	12.2	*100.0*

Note: Population considered are single-job workers age 15 and above, working 30 and more hours weekly

Unfortunately, it is not possible to quantify the precise effect of changes in the returns to unobserved skill on the formal-informal wage gap, due to the fact that the distribution of unobserved skill between formals and informal workers is unknown and may have changed over time. This is particularly likely given the large expansion in informal work over time. Note that the returns to unobserved skill analysis presented in the previous section relies on the plausible assumption that within skill (education and experience) groups, the distribution of unobserved skill has not changed over time. Because of the potential for workers to move between formal and informal jobs, a similar assumption for formal and informal worker groups would not be tenable.

If, as these results suggest, the growing informal-formal gap has been driven by increases in the returns to unobserved skill, technical change may be the ultimate cause of some of the informal-formal dynamics. Changes in technology employed in formal salaried jobs may have increased the demand for workers with complementary skills. If these skills are unobserved in survey data (but observed by potential employers), those with unobserved skill may have been sorted into formal sector jobs, expanding the gap between informal and formal jobs, even while the share of informal employment has increased.

5.7 Conclusion

The variance decomposition analysis presented in this paper shows that that the returns to unobserved skill decreased in Argentina in the 1980s and then increased during the 1990s, during the same period that the returns to education increased. The changes in the 1980s are compatible with a stagnation of the level of technology in Argentina during the decade, paired with educational upgrading of the workforce. In other words, the drop in returns may reflect an increasing supply of those skills, combined with falling demand.

We interpret the simultaneous timing of the increase in the returns to unobserved skill and education in the 1990s as circumstantial evidence that skill-biased technical change may be driving both phenomena. Demand for relevant skills outpaced the parallel increase in supply of those skills in the workforce with improved education.

The variance decomposition also demonstrated that in the 1990s the growth in the inequality of the upper half of the earnings distribution was mainly caused by rising returns to college education. In contrast, the growth in the inequality of the lower half of the distribution can mainly be explained by increasing returns to unobserved skills. This finding offers a possible explanation for the growth in the wage gap between informal and formal salaried workers in the 1990s. Competing explanations to SBTC do not seem to be able to explain these phenomena. Also, changing returns to education do not account for the changes in relative wages. As the increase in the returns to unobserved skill has taken place largely between the median and the lower end of the distribution, where informal workers are concentrated, this could be interpreted as evidence that changes in the returns to unobserved skill have driven the relative drop in the wages of informal workers.

References

Acemoglu, Daron (2002). Technical Change, Inequality, and the Labor Market. *Journal of Economic Literature, 40*(1), pp. 7–72.

Acosta, Pablo, Gasparini, Leonardo (2007). Capital Accumulation, Trade Liberalization, and Rising Wage Inequality: The Case of Argentina. *Economic Development and Cultural Change, 55*(4), pp. 793–812.

Blau, Francine D., Kahn, Lawrence M. (1997). Swimming Upstream: Trends in the Gender Wage Differential in 1980s. *Journal of Labor Economics, 15*(1), pp. 1–42.

Card, David (1999). Chapter 30 The Causal Effect of Education on Earnings. In Orley C. Ashenfelter and David Card (Ed.), *Handbook of Labor Economics*. Volume 3, Part 1, pp. 1801–1863. Elsevier.

Card, David, Dinardo, John (2005). The Impact of Technological Change on Low Wage Workers: A Review. National Poverty Center Working Paper Series No. 05-28.

Chay, Kenneth Y., Lee, David S. (2000). Changes in Relative Wages in the 1980s: Returns to Observed and Unobserved Skills and Black-White Wage Differentials. *Journal of Econometrics, 99*(1), pp. 1–38.

Cicowiez, Martín (2003). Comercio y Desigualdad Salarial en Argentina: Un Enfoque de Equilibrio General Computado. CEDLAS Documento de Trabajo No. 3.

Demombynes, Gabriel, Metzler, Johannes (2008). Connecting the Unobserved Dots: A Decomposition Analysis of Earnings Inequality Changes in Urban Argentina, 1980-2002. World Bank Policy Research Working Paper No. 4624.

Dinardo, John, Fortin, Nicole M., Lemieux, Thomas (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica, 64*(5), pp. 1001–1044.

Galiani, Sebastian, Sanguinetti, Pablo (2003). The Impact of Trade Liberalization on Wage Inequality: Evidence from Argentina. *Journal of Development Economics, 72*(2), pp. 497–513.

Gasparini, Leonardo C. (2002). Microeconometric Decompositions of Aggregate Variables: An Application to Labour Informality in Argentina. *Applied Economics, 34*(18), pp. 2257–2266.

Gasparini, Leonardo C. (2003). Argentina's Distributional Failure: The Role of Integration and Public Policies. CEDLAS Documento de Trabajo No. 1.

Giovagnoli, Paula I., Fiszbein, Ariel, Patrinos, Harry A. (2005). Estimating the Returns to Education in Argentina: 1992-2002. World Bank Policy Research Working Paper No. 3715.

Goni, Edwin, Maloney, WIlliam F. (2007). *On the Effects of Trade on Informality. A Note for the Argentine Case.* Unpublished manuscript, World Bank.

Juhn, Chinhui, Murphy, Kevin M., Pierce, Brooks (1993). Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy, 101*(3), pp. 410–442.

Katz, Lawrence F., Murphy, Kevin M. (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *Quarterly Journal of Economics, 107*(1), pp. 35–78.

Khamis, Melanie (2007). *Does the Minimum Wage Have a Higher Impact on the Informal Than on the Formal Labor Market? Evidence from Quasi-experiments.* Unpublished manuscript.

Lemieux, Thomas (2002). Decomposing Changes in Wage Distributions: A Unified Approach. *Canadian Journal of Economics, 35*(4), pp. 646–688.

Lemieux, Thomas (2006). Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill? *American Economic Review, 96*(3), pp. 461-461.

Maloney, WIlliam F., Mendez, Jairo N. (2003). Measuring the Impact of Minimum Wages: Evidence from Latin America. National Bureau of Economic Research Working Paper No. 9800.

Margot, Diego (2001). Rendimiento a la education en Argentina: Un análisis dinámico basado en cohortes. UNLP Working Paper.

Marshall, Adriana (2005). Labor Regulations and Unionization Trends: Comparative Analysis of Latin American Countries. Visiting Fellow Working Papers. Cornell University.

Mincer, Jacob (1974). *Schooling, Experience, and Earnings* (Vol. 2). New York: National Bureau of Economic Research.

Porto, Guido (2003). Trade Reforms, Market Access and Poverty in Argentina. World Bank Policy Research Working Paper No. 3135.

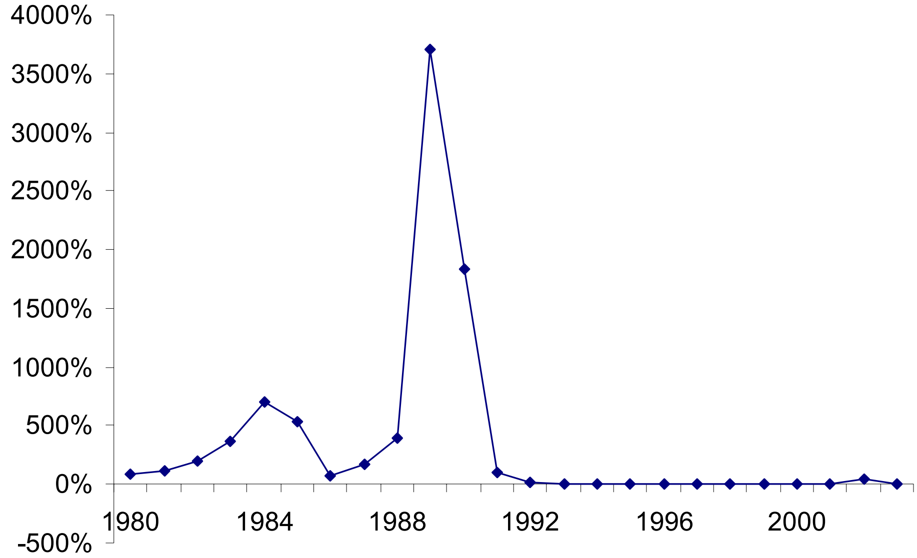
World Bank (2003). Closing the Gap in Education and Technology. World Bank Latin American and Caribbean Studies. Washington, D.C.: World Bank.

World Bank (2008). *Informal Employment in Argentina: Causes and Consequences.* Washington, D.C.: World Bank.

Appendix 5

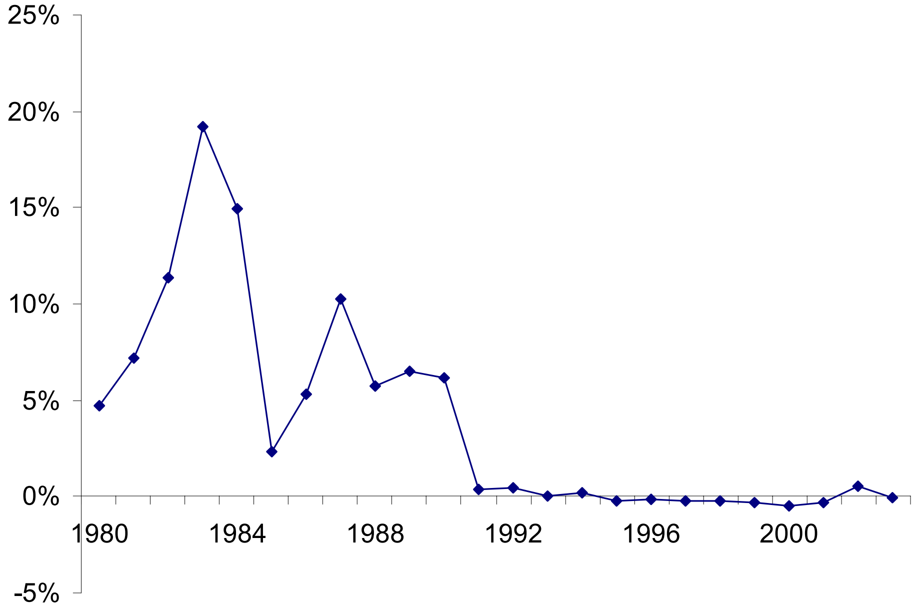
Appendix 5.1 Inflation in Argentina

Figure 5-6. Yearly Inflation in Buenos Aires, 1980-2003



Source: INDEC CPI

Figure 5-7. October Inflation in Buenos Aires, 1980-2003



Source: INDEC CPI

Appendix 5.2 Extended Decomposition Results

Table 5-14. Decomposition Results, 1980-2002

Panel A. Men

	Var logwage	90/10	50/10	90/50
2002	0.684	7.882	2.361	3.338
1980	0.481	5.488	2.091	2.625
Change	0.203	2.394	0.271	0.713
Components of change:				
Δ Returns to Observed Skills	0.099	1.201	0.071	0.421
Δ Composition of Observed Skills	0.053	1.101	-0.095	0.578
Δ Returns to Unobserved Skills	0.051	0.092	0.295	-0.286

Panel B. Women

	Var logwage	90/10	50/10	90/50
2002	0.649	7.361	2.577	2.856
1980	0.465	5.065	2.101	2.411
Change	0.183	2.296	0.476	0.445
Components of change:				
Δ Returns to Observed Skills	0.082	1.710	0.058	0.613
Δ Composition of Observed Skills	0.049	-0.181	-0.349	0.209
Δ Returns to Unobserved Skills	0.052	0.768	0.767	-0.377

Note: Estimations for single-job workers age 15 or above working 30 or more hours weekly

Table 5-15. Decomposition Results, 1992-2002, ARG16 Sample

Panel A. Men

	Var logwage	90/10	50/10	90/50
2002	0.661	7.598	2.495	3.045
1992	0.457	5.224	1.983	2.635
Change	0.204	2.374	0.513	0.410
Components of change:				
Δ Returns to Observed Skills	0.092	1.248	0.092	0.402
Δ Composition of Observed Skills	0.024	0.482	0.057	0.141
Δ Returns to Unobserved Skills	0.088	0.644	0.364	-0.134

Panel B. Women

	Var logwage	90/10	50/10	90/50
2002	0.658	7.576	2.840	2.668
1992	0.374	4.436	1.940	2.286
Change	0.284	3.141	0.900	0.382
Components of change:				
Δ Returns to Observed Skills	0.136	1.799	0.139	0.529
Δ Composition of Observed Skills	0.007	-0.115	-0.070	0.012
Δ Returns to Unobserved Skills	0.141	1.456	0.830	-0.159

Note: Estimations for single-job workers age 15 or above working 30 or more hours weekly

Curriculum Vitae
Johannes Moritz Metzler

10/05– 09/08	Ph.D. Program in Economics Ludwig Maximilians-Universität, Munich Munich Graduate School of Economics
08/04 – 05/05	Advanced Studies Program in International Economic Policy Research Institut für Weltwirtschaft, Kiel
09/00 – 06/04	Diplom-Kaufmann (M.B.A. equivalent) Studies in Business Administration European Business School, Oestrich-Winkel
01/03 – 05/03	Exchange Student San Diego State University, San Diego, USA
07/02 – 12/02	Exchange Student Universidad Argentina de la Empresa, Buenos Aires, RA
09/91 – 06/00	Abitur Kurfürst-Friedrich-Gymnasium, Heidelberg
6 April 1981	born in Heidelberg