UNIVERSITÉ DE SHERBROOKE

École de gestion

CHOCS SUR LE MARCHÉ DU TRAVAIL ET DÉCISIONS D'ADAPTATION DES MÉNAGES EN INDE

PAR

BOUBACAR DIOP

Thèse présentée à l'École de gestion en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.) en économie du développement

12 AVRIL 2022

École de gestion

CHOCS SUR LE MARCHÉ DU TRAVAIL ET DÉCISIONS D'ADAPTATION DES MÉNAGES EN INDE

PAR

BOUBACAR DIOP

Thèse présentée à l'École de gestion en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.) en économie du développement

Cette thèse a été évaluée par un jury composé des personnes suivantes :

P^r Patrick Richard, Ph.D	Président du jury et examinateur interne
P^r Martino Pelli, Ph.D	Directeur de thèse
P^r Jean-François Rouillard, Ph.D	Examinateur interne
P^r Beyza Ural Marchand, Ph.D	Examinatrice externe

SOMMAIRE

Ma thèse étudie les décisions des ménages ruraux en Inde concernant leur participation au marché du travail suite à un choc positif (productivité agricole) ou négatif (catastrophe naturelle). Les principales hypothèses qui seront testées dépendront de la nature du choc auquel le ménage est confronté.

Le premier article se concentre sur un choc de productivité positif résultant de l'adoption d'une technologie agricole (électricité pour l'irrigation). Dans ce cas, le ménage peut décider de libérer la main-d'oeuvre familiale des tâches agricoles et de permettre aux individus de s'engager dans d'autres industries pour diversifier les sources de revenus du ménage. Si une telle hypothèse est vraie, alors l'agrégation de ces comportements individuels conduirait à une transformation structurelle de l'économie indienne. Si, en revanche, le ménage décide de profiter de ce choc de productivité pour augmenter ses revenus agricoles en augmentant la surface cultivée et / ou la main-d'oeuvre employée, il est légitime de s'attendre à une pression sur les ressources naturelles telles que la forêt ou la nappe phréatique. Pour ces raisons, il est essentiel de déterminer comment l'adoption de la technologie agricole affecte les facteurs travail et terre dans le processus de production agricole.

Le deuxième article étudie comment la politique commerciale influence la relation entre la productivité agricole et la transformation structurelle des pays en développement. Théoriquement, il est prouvé qu'une augmentation de la productivité agricole peut conduire à une industrialisation ou à une spécialisation dans le secteur agricole, selon que l'on se trouve dans une économie fermée ou ouverte. Je profite de la réforme commerciale de l'Inde de 1991 et de l'utilisation de l'électricité à des fins d'irrigation dans le secteur agricole pour étudier empiriquement cette question. Je constate une diminution de la part de la main-d'oeuvre agricole dans les districts qui ont connu une plus forte baisse du niveau tarifaire et une augmentation de la part de la main-d'oeuvre dans les autres secteurs de l'économie.

Le troisième article de cette thèse se concentre sur les conséquences d'un choc négatif de catastrophe naturelle sur les décisions d'allocation de la main-d'oeuvre au sein du ménage. En effet, les décisions prises par les parents concernant l'éducation de leurs enfants auront des répercussions futures sur leur participation au marché du travail. Un ménage confronté à un choc de revenu négatif peut décider d'interrompre l'éducation de ses enfants et de les utiliser comme main-d'oeuvre supplémentaire pour augmenter le revenu familial. L'idée est donc de montrer comment les cyclones affectent l'éducation des enfants en Inde à court ou moyen terme et quel sera l'impact à long terme sur leur participation au marché du travail.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

REMERCIEMENTS

Je remercie chaleureusement toutes les personnes qui m'ont aidé pendant l'élaboration de ma thèse et notamment mon superviseur Pr. Martino Pelli, pour son intérêt et son soutien, sa grande disponibilité et ses nombreux conseils durant la rédaction de ma thèse. Mes remerciements également à mon comité de thèse, Pr. Patrick Richard, Pr. Jean-François Rouillard et Pr. Beyza Ural Marchand pour leurs commentaires et suggestions qui vont m'aider à améliorer mon travail et à m'améliorer en tant que chercheur. Ce travail n'aurait pas pu être mené à bien sans la disponibilité et le soutien de mes camarades doctorants de l'École de gestion de l'Université de Sherbrooke. Je remercie enfin celles et ceux qui me sont très chers et qui m'ont accompagnés durant ces années difficiles. Je suis infiniment redevable à mes parents pour leur soutien moral et matériel et leur confiance indéfectible dans mes choix. Je ne peux pas terminer sans remercier mon épouse bien-aimée avec qui je partage chaque moment.

CHAPITRE 1
INTRODUCTION GÉNÉRALE

Les économistes du développement ont pour principal objectif de proposer des politiques économiques en vue d'améliorer le bien-être et les conditions de vie des individus dans les pays en développement. Ce bien-être est en partie dépendant des interactions économiques et sociales qu'ont les individus avec les institutions dans lesquelles ils sont intégrés qu'est le ménage ou la communauté. De nombreux gouvernements ou Organisations Non Gouvernementales (ONG) assument l'hypothèse selon laquelle le bien-être individuel peut s'accroître en améliorant le bien-être collectif du ménage ou de la communauté. D'autres, par contre, mettent en place leurs politiques en dissociant bien-être individuel et collectif et soutiennent que le premier peut être amélioré sans tenir compte du second. Dès lors, il est important de bien identifier l'influence de la communauté sur l'individu afin de mettre en place des politiques économiques efficaces.

Becker [1962] est l'un des premiers économistes à placer le ménage au coeur des décisions d'allocation des ressources en maximisant une fonction d'utlité au niveau ménage où ce dernier agit comme une seule et unique entité. Plusieurs travaux ont par la suite émergés dans la littérature, en considérant la prise de décisions au sein du ménage comme un processus unitaire, avec une réallocation des ressources qui refléte les décisions des/du chef(s) de ménage ; ou bien comme un processus collectif où les avis des différents membres du ménage sont pris en compte [Ashraf, 2009, Ashraf, Field, and Lee, 2014, Attanasio and Lechene, 2002, Banerjee, 2004, Browning and Chiappori, 1998, Hoddinott and Haddad, 1995, Pitt, Rosenzweig, and Hassan, 1990, Thomas, 1990, Udry, 1996].

Ma thèse s'insère dans cette littérature et examine comment est-ce qu'un choc positif ou négatif sur les ressources du ménage peut affecter les décisions de réallocation de la main-d'oeuvre ou de formation du capital humain au sein de la famille. Et quelles peuvent être les conséquences sur l'économie dans sa globalité lorsque l'on agrège ses comportements individuels.

Le chapitre 1 considère un choc positif de productivité agricole et étudie comment est-ce qu'il affecte la réallocation de la main-d'oeuvre entre les secteurs de l'économie ainsi que l'exploitation des ressources naturelles. De manière spécifique, il considère l'utilisation de l'électricité à des fins d'irrigation comme une technologie capable d'augmenter la productivité des petits exploitants agricoles en Inde. Deux hypothèses principales sont testées. La première considère que les ménages profitent de la disponibilité de la technologie pour diversifier leurs sources de revenus. En effet, l'utilisation de l'électricité à des fins d'irrigation peut accroître la productivité du facteur travail, puisque la pompe électrique peut réaliser le travail de plusieurs individus. Ainsi, le surplus de la main-d'oeuvre familliale peut être réaffecté à d'autres secteurs de l'économie pour permettre une diversification des sources de revenu du ménage. Une agrégation de ces comportements individuels peut mener à une transformation structurelle de l'économie. La deuxième hypothèse considère plutôt que les ménages agricoles profitent de l'utilisation de l'électricité pour accroître leur revenu agricole, en augmentant leurs surfaces cultivées ou en adoptant des cultures plus intensives en eau. Si une telle hypothèse se vérifie, on peut donc s'attendre à une déforestation pour augmenter les terres agricoles ou une surutilisation de la nappe phréatique pour arroser une plus grande surface ou des types de cultures plus intensives en eau. Ce chapitre contribue à deux types de littératures : celle qui lie la productivité agricole à la transformation structurelle des pays en développement et celle qui lie la productivité agricole à l'exploitation des ressources naturelles.

Le chapitre 2 de cette thèse est lié au premier et examine empiriquement l'effet de l'ouverture commerciale sur la relation entre la productivité agricole et la transformation structurelle des pays en développement. En effet, la litérature a théoriquement établit que l'effet de la productivité agricole sur la transformation structurelle des économies dépend fortement de l'hypothèse d'économie ouverte ou fermée qui est mis en avant. Cependant, il est difficile de vérifier cette hypothèse empiriquement parce qu'il n'existe au monde aucun pays complétement ouvert ou fermé au commerce mondial. En utilisant la réforme commerciale de l'Inde de 1991 ainsi qu'une mesure des tarifs spécifique au district inspiré d'Edmonds, Pavcnik, and Topalova [2010a], le chapitre 2 arrive à vérifier empiriquement comment est-ce qu'une politique commerciale peut affecter le lien entre productivité agricole et transformation structurelle d'un pays en développement. Ce chapitre contribue à la litérature à deux niveaux. D'abord, il étudie la transformation structurelle des districts, considérés comme de petites économies ouvertes, soumis au même choc et au même moment. Alors que la littérature compare souvent la transformation structurelle de pays avec des caratéristiques socio-économiques différentes et souvent à différentes périodes. Ensuite, ce chapitre propose une approche par variable instrumentale pour essayer d'établir une relation causale de l'effet de l'ouverture commerciale sur la relation entre productivité agricole et transformation structurelle.

Le troisième et dernier chapitre de cette thèse considère l'effet d'un choc négatif de désastre naturel sur l'éducation et l'emploi des individus à long terme. Plus précisément, ce chapitre utilise les cyclones qui se produisent durant les années de scolarité obligatoire des individus en Inde pour étudier ses effets sur le retard scolaire des individus ainsi que le type d'emploi qu'ils occupent lorsqu'ils deviennent de jeunes adultes. En effet, les désatres naturels peuvent affecter l'éducation des individus à travers deux canaux : l'offre et la demande d'éducation. Du côté de l'offre, les cyclones peuvent causer l'interruption temporaire ou définitive de l'éducation des individus, en détruisant notamment les infrastructures scolaires et routières. Du côté de la demande, les effets se manifestent au niveau de la santé mentale des individus ou à travers un choc négatif sur le revenu des ménages. L'offre et la demande d'éducation sont aussi les deux canaux par lesquels la COVID-19 affecte l'éducation des individus avec la fermeture obligatoire des écoles imposée par les gouvernements, les troubles mentaux générés par la pandémie ainsi que le choc négatif sur le revenu des individus lié aux pertes massives d'emploi. C'est la raison pour laquelle, ce chapitre contribue, dans une moindre mesure, à la littérature grandissante sur la COVID-19 et peut servir de base pour étudier les impacts à long terme de la pandémie. Il contribue également à deux autres types de littérature. La première est liée à l'éducation dans les pays en développement et la seconde est liée à la relation entre éducation et marché du travail.

Ainsi, en considérant un choc positif dans les chapitres 1 et 2 et un choc négatif dans le chapitre 3, cette thèse montre les décisions prises par les ménages en Inde en terme d'utilisation de la main-d'oeuvre ou de formation du capital humain au sein du ménage.

CHAPITRE 2

STRUCTURAL TRANSFORMATION OR LOSS OF NATURAL RESOURCES : AN INDIAN DILEMMA

1 AVANT-PROPOS

Je suis le seul auteur sur cet article. Le projet de recherche, le traitement de la base de données, la production des résultats ainsi que la rédaction sont le fruit de mon travail, sous la supervision de Pr. Pelli.

2 RÉSUMÉ

J'étudie l'effet de l'utilisation de l'électricité pour l'irrigation sur la transformation structurelle de l'économie indienne et ses répercussions sur les ressources naturelles. Dans le travail empirique, j'aborde les problèmes d'endogénéité de l'électrification en utilisant les variations de l'offre au niveau de l'état dans la disponibilité de l'énergie hydroélectrique pondérée par le niveau initial d'électrification de chaque district (1992) comme variable instrumentale. Les résultats suggèrent que l'adoption de l'électricité pour l'irrigation permet de réduire la quantité de main-d'oeuvre et de terre car elle augmente les rendements agricoles, avec moins de facteurs de production requis. J'observe une réallocation de la main-d'oeuvre du secteur agricole vers les secteurs non-agricoles avec une analyse au niveau du district. Ceci est confirmé par l'analyse au niveau ménage, avec une diminution des membres du ménage travaillant dans l'agriculture et une augmentation dans le secteur non-agricole. Je ne trouve aucun effet significatif sur l'épuisement de la nappe phréatique. Enfin, l'électrification a un effet négatif statistiquement significatif sur la couverture forestière.

3 ABSTRACT

I study the effects of electricity use for irrigation on the structural transformation of the Indian economy and its repercussions on natural resources. In the empirical work, I address the endogeneity problems of the electricity variable by using the state-level supply shifts in hydro-electric power availability weighted by the initial level of electrification of each district (1992) as an instrumental variable. The findings suggest that electricity adoption for irrigation is labor and land-saving as it increases agricultural yields, with less factor of production required. I observe a labor reallocation from the agricultural sector towards the manufacturing sector with an analysis at the district level. This is confirmed by the analysis at the household level, with a decrease of household members working in agriculture and an increase in non-agriculture. I find no significant effect on water table depletion. Finally, electrification have a statistically significant negative effect on forest cover.

4 INTRODUCTION

The recent literature on the adoption of new agricultural technologies can be divided into two parts. First, researchers focus on the fact that the development of the agricultural sector is at the origin of the development process of countries. In particular, an increase in agricultural productivity is an essential condition for the structural transformation of an economy because of the reallocation of physical and human capital between sectors that it generates [Bustos, Caprettini, and Ponticelli, 2016, Emerick, 2018, Foster and Rosenzweig, 2004, 2007, Hornbeck and Keskin, 2015, Rud, 2012]. Second, some authors question the consequences of the use of new agricultural technologies on natural resources. According to the World Bank (2012), an increase in agricultural productivity is the best way to preserve natural resources from an uncontrolled use. However, it is important to consider farmers behaviour when they have at their disposal a technology that enhances their productivity and profits. One can expect them to want to cultivate more land since their activity becomes more profitable. An uncontrolled use can have repercussions on the preservation of the environment, such as an increase in deforestation or a decrease of the water table [Angelsen and Kaimowitz, 2001, Fishman, Devineni, and Raman, 2015].

The Green Revolution in India, which began around 1967/1968, is a perfect occasion to measure the effect of a technology adoption in the agricultural sector. After the Bengal Famine of 1943, the Indian government's desire was to increase agricultural productivity by adopting various technologies. Despite the achievement of the objective of food self-sufficiency, it has had negative consequences on the environment, mainly with soil, vegetation and water resources degradation [Singh, 2000]. One component of this policy is the supply of electricity to small-scale farmers in order to meet their irrigation needs. That is why the Indian government had put in place various policies and programs ranging from the Public Distribution System (PDS) of kerosene in 1957 to the Rural Electrification Policy in 2006 [Balachandra, 2011]. Electricity supply to farmers was therefore of central importance and became a campaign tool for politicians, so that by the mid-1980's, power consumption in the agricultural sector was well above the one in the industrial sector [Dubash, 2007]. How important are the effects of electricity for irrigation on the labor reallocation between sectors in rural India ? Does it favour deforestation or the depletion of the water table ?

As pointed out in the literature by Matsuyama [1992] and Bustos et al. [2016], the effect of agricultural productivity on industrial development depends on several factors such as the degree of openness of an economy. Using a closed economy assumption, Matsuyama [1992] shows that an agricultural revolution is a precondition for an industrial revolution. In fact, the use of an agricultural technology increases the output per worker and output per unit of land. The surplus of workers in the agricultural sector is then reallocated towards the manufacturing sector, leading to its expansion. In contrast, the more realistic assumption of open economy proves a reallocation of the workforce towards the agricultural sector. In accord with the theory of comparative advantage, the increase in agricultural productivity pushes the domestic economy to specialize in the production of the agricultural good to the detriment of the non-agricultural good, leading to a progressive de-industrialization of the economy. Bustos et al. [2016] come to the same conclusions by considering how intensively the use of an agricultural technology affects the use of various inputs. A labor-saving technology, which is supposed to decrease labor input per unit of land, may lead to economic growth through the reallocation of part of the workforce towards other sectors. A labor-biased technology, which has the particularity to increase labor input requirement, may dampen industrial expansion.

The consequences of electricity use for irrigation purposes in India might be numerous, depending on many factors such as farmers behaviour. If electricity is for example a labor-saving technology, one can expect a smallholder farmer employing family labor to release household members from farm tasks. This should enable the household to diversify its source of income, with the surplus of family labor moving towards other sectors of activity. This reorganization of the workforce at the household level may lead to a labor supply shock in other sectors and other areas. Gollin, Lagakos, and Waugh [2013] show that in developing countries, value added per worker is four times higher in non-agricultural sector than in agriculture while the labor share in the agricultural sector is much higher, attesting a misallocation of labor between sectors. In the case of India, the labor share in the agricultural sector was around 68.1% in the 1980s [Kochhar, Kumar, Rajan, Subramanian, and Tokatlidis, 2006], and declined over the years to 48.9% in 2011-12 (International Labour Organization, 2016). But India is known to have an atypical structural transformation, as most of the agricultural labor is absorbed by the services sector. According to Manisha Goel and Paulina Restrepo-Echavarria (2015) the manufacturing labor share has increased faster than the services labor share since 2000. [1] This point is an additional argument that justifies

1. https://research.stlouisfed.org/publications/economic-synopses/2015/10/02/indias-atypical-structural-transformation/

paying attention to the cross-sectoral reallocation of the agricultural workforce and which sector absorbed most of this labor. De Janvry, Emerick, Gonzalez-Navarro, and Sadoulet [2015] show that a large-scale adjustment of labor may induce the migration of the workforce and drive down wages in other sectors. In this case it will be interesting to see what may be the consequences of the resulting increase of labor supply in terms of wage level in the manufacturing sector.

Instead of releasing family labor, another possibility for farmers would be to increase the cultivated area in order to increase the farming income of the family. Fishman et al. [2015] describe this assumption of increasing cultivated area as a realistic scenario where farmers act rationally and want to take advantage of the availability of the new technology to maximize their income. In the same line of thought, Angelsen and Kaimowitz [2001] argue that farmers would certainly want to cultivate more land if a profitable technology that increases their yields or lowers their input costs is available. This situation would increase the demand of arable land and of water to irrigate larger surfaces. If electricity used for farming activities in India increases yields or lowers agricultural costs, one can legitimately ask the question of how it influences the use of environmental resources. The same situation can be expected if electricity is a labor-biased technology. In this case, an influx of labor to the agricultural sector would also accentuate the demand of arable land and water for irrigation, accentuating the pressure on the forest cover or on the water table.

The main challenge when studying this subject is that electricity is not exogenous to agricultural and manufacturing outcomes. For instance, there might be unobservable (political/economic) variables that are correlated to the level of electrification of some areas compared to others. In addition, the three sources of endogeneity - measurement error, reverse causality and omitted variable - may be valid. First, I use the night lights, captured from satellite images, as a measure of the level of electrification of rural areas and I consider it as a proxy of the quantity of electricity available for irrigation, raising a potential problem of measurement error that cause the endogeneity of the electricity variable. Second, reverse causality can also be established between agricultural land expansion and the use of electricity for irrigation. For example, an increase of the area cultivated would require more sophisticated system of irrigation such as drip, sprinkler or pump that use electricity. The link can also be established in the opposite way. A rational farmer who increases its Total Factor Productivity (TFP) by using electricity for irrigation would probably like to expand its land, in order to increase his earnings. Finally, the third threat of endogeneity is related to omitted variables. For example, farmers with better entrepreneurial skills would be the ones who adopt electricity for irrigation and increase their earnings. Such skills that are difficult to measure will be therefore in the error term and correlated with electricity adoption.

I use the instrumental variable developed by Allcott, Collard-Wexler, and O'Connell [2016a] to takle this endogeneity problem and establish empirical evidence. In their paper, authors use the state level supply shifts in hydro-electric power to instrument electric shortages and show its effects on the performance of manufacturing plants in India. I then use a Bartik instrument that is the interaction of the state level instrument with the initial level of electrification of each district.[2] This strategy allows to transform the state level instrument to a district level one.

I, first, look at the effect of electricity for irrigation on the labor and land factor in the agricultural production process. The outcomes of interest are the area cultivated, the agricultural labor share, the labor intensity and the agricultural production. This analysis

2. Several other papers used the same strategy in the literature : Bartik [1991], Blanchard, Katz, Hall, and Eichengreen [1992], Bound and Holzer [2000], Autor and Duggan [2003], Chakravorty, Pelli, and Ural [2014], Gupta and Pelli [2021]. See Goldsmith-Pinkham, Sorkin, and Swift [2020] for the latest development on the Bartik instrument.

will allow to know if there is a reduction (factor-saving) or an increase (factor-biased) of labor and/or land required for agricultural production, following the use of electricity. I, then, look at the repercussions on the manufacturing sector and on natural resources.

Using data from rural India, my results show that an increase of electricity used for irrigation significantly affects labor reallocation between the agricultural and the manufacturing sectors. In fact, a one unit increase in night lights can be associated with a decrease of the agricultural labor share by 2.2% and an increase for the manufacturing other non-agricultural sectors by 0.8% and 1.5% respectively. This finding is confirmed by the household level analysis with an increase of the non-agricultural workers by 0.011 individuals and a decrease of the agricultural workers within the household by 0.012 individuals, but the coefficient is not statistically significant for agricultural workers. I find also that areas with faster adoption of electricity for irrigation experience a reduction of the area cultivated by 4.7%. The analysis on the consequences on natural resources suggests that there is not a direct effect of electricity for irrigation on forest cover. And there is no significant effect on the depletion of the water table.

The following section describes the literature related to the effect of agricultural technology adoption. Section 6 presents an overview of the rural electrification issues in India and show how important is electricity for both farmers, in their farming activities, and politicians, to attract the votes of rural citizens. Section 7 presents the different sources of data that I use. As I use the night lights emission as a proxy of the quantity of electricity available in a district, I show its correlations with the area equipped with an irrigation system in section 8. In section 9 I present the empirical strategy that I use, while section 10 presents the results. I investigate some potential mechanisms through which electricity for irrigation affects the Indian economy in section 11. And section 12 concludes.

Understanding the role of agriculture in the development process of countries is one of the main tasks pursued by researchers. Earlier works of Rosenstein-Rodan [1943], by using the principle of international division of labor and Nurkse [1953], who studied the role of agriculture in a densely populated and sparsely populated countries, laid the foundation of what will become a central objective of development economics.

A vast body of research has theoretically examined the link between agricultural productivity and industrial development through two distinct assumptions of closed and open economies. The first group of economist who made their analysis in closed economy assumption present two ways through which agricultural productivity can favour industrial development : the demand and the supply channel [Gollin, Parente, and Rogerson, 2002, Kongsamut, Rebelo, and Xie, 2001, Murphy, Shleifer, and Vishny, 1989]. The second group, with the open economy assumption, argues that using a new technology in the agricultural sector can lower industrial development because of the theory of comparative advantage. In fact, an increase in agricultural productivity may increase the comparative advantage of the agricultural sector, leading to a specialization in this sector [Field, 1978, Matsuyama, 1992, Wright, 1979]. More recent papers in the field develop a theoretical macroeconomic framework that allow to compare a situation of autarky or free trade, by calibrating the model on the economy of developing countries [Echevarria, 1997, Gollin, Parente, and Rogerson, 2007, Hayashi and Prescott, 2008, Teignier, 2018].[3] For instance, Teignier [2018] compare the structural transformation path of the Great Britain during the industrial revolution with South Korea for the period 1963-2015. He mainly concludes that developing countries should

3. See Herrendorf, Rogerson, and Valentinyi [2014] and Gabardo, Pereima, and Einloft [2017] for a comprehensive review of this literature.

open their economy to international trade. This strategy will allow them to specialize in the production of the good where they have a better productivity and to import from the rest of the world other goods that they can not produce.

This paper is mainly related to the empirical literature that study the effect of agricultural productivity increase and structural transformation in developing countries. Foster and Rosenzweig [2008, 2004] make a considerable contribution to the literature by studying the effects of High Yielding Variety (HYV) seeds during the Green Revolution in India. They mainly find that areas that have a faster increase of agricultural productivity experience a lower development of the manufacturing sector. Several other researchers contribute to the literature of structural transformation of the Indian economy [Asher, Campion, Gollin, and Novosad, 2021, Emerick, 2018, Rud, 2012]. Bustos et al. [2016] explored factor-biased and factor-saving technologies adoption in the agricultural sector in Brazil. The authors showed that there is a development of the non-agricultural sector when technology is labor-saving and labor is reallocated towards the agricultural sector in the case of a labor-biased technology. A similar approach has been adopted by Hornbeck and Keskin [2015], to show the effect of using improved pumps to reach the Ogallala aquifer in USA on local economic spillovers. They base their empirical strategy on the fact that counties over the Ogallala were similar to counties in the same state prior to 1940. Thus, without the possibility to access the aquifer, these counties would have changed similarly. Their main result is that agricultural productivity may not generate local economic spillovers due to, mainly, an increase of the labor cost in non-agricultural sector.

This work also builds on the literature related to the relationship between agricultureal productivity and deforestation. In their book, Angelsen and Kaimowitz [2001] present a review of economic models of deforestation. One of them is the work of Cattaneo (2001) who uses a Computable General Equilibrium (CGE) model to analyse the link between technological change in agriculture, migration and deforestation in the Brazilian Amazon. The author justifies his methodology by the fact that forces underlying deforestation process are multiple and range from macroeconomic policies to biophysical processes that change land cover. He argues that in the long-run, agricultural technological change in a region allows to attract economic resources (labor and capital) from other regions or other economic activities. And it is legitimate to think that this fact may increase the demand of arable land, which then leads to an increase of deforestation. Kaimowitz and Smith [2001] study the effect of soybean technology on loss of natural vegetation in Brazil and Bolivia and the general equilibrium effects it generated in labor and product markets. They point out that the structural transformation of the economy may be triggered by economies of scale at a sector level. In fact, they say that "technical change can make it easier to profitably reach level of production that justifies installing ancillary services and infrastructures". This contribute to improve the business environment favourable to the development of other business activities.

Finally, this paper is also related to the relationship between agricultural productivity and groundwater extraction in India. Fishman et al. [2015] give some thoughts on how to preserve the water table in India. They report that irrigated agriculture threatens water resources in developing countries because of abusive exploitation. The study they conducted led them to find that the widespread adoption of ground pumping technologies such as drip and sprinkler, that require electricity, may allow to reduce the amount of excessive extraction of groundwater by two thirds. The effect of electricity subsidies on the water table exploitation in India has been documented by Badiani and Jessoe [2013]. They argue that even if the use of electricity in farming activities raises notably agricultural productivity, it harms the environment with an over-exploitation of the groundwater.

6 OVERVIEW OF RURAL ELECTRIFICATION ISSUES IN INDIA

In this section, I present some background information on rural electrification in India. The country was characterized by a low agricultural productivity during the post-colonial period, that led to the Bengal-famine in 1943.[4] After the independence in 1947, one of the main priorities of the Indian government was to overcome this food shortage, by increasing agricultural productivity. That is why the Green Revolution was launched during the period 1967/68 with the objective of achieving food self-sufficiency.

The Green Revolution had several components, including land reform, irrigation infrastructure installation, herbicide and pesticide use, etc. One component of the program was to provide electricity to farmers for their irrigation needs. For this reason, in each state a State Electricity Board (SEB) was created, with the Electricity Act of 1948. They were responsible of the transmission, distribution and generation of electricity, as well as to support rural electrification schemes and rural electricity co-operatives [Badiani and Jessoe, 2013]. The creation of SEBs allowed public control on the management of the power sector in India. Despite these efforts by the government to achieve universal access to electricity, the objectives were difficult to reach mainly because of the costs they generated. According to Lamb [2005], the financial losses of the SEBs reached US$ 6 billion, representing 1.3 % of India's GDP. With the intention of reducing this financial burden and being more effective in achieving the goals, the government reformed the power sector in the early 1990's in order to reduce the public sector control and to let private initiatives meet the shortfall in generation capacity. As a result, the Electricity Laws Act of 1991 gives to Independent Private Providers (IPPs) the possibility to operate in the power sector. India's government considers that the

4. According to Sen, Hobsbawm, et al. [1980], around 2.7 million individuals died of starvation during that famine, due in large part to poor agricultural performance.

consequences of this reform are important on rural electrification. In fact, it is estimated that 481,124 out of 634,321 villages were electrified by 1991, while only 1500 were able to access to electricity in 1947. [5]

According to Modi [2005], rural electrification in India also covers political issues that politicians use in their political constituencies. In addition, Badiani and Jessoe [2013] assert that there is a fierce political competition among politicians to attract the vote of farmers who express the need to have stable water supply for agricultural purposes. Electricity subsidy was therefore used as a campaign tool, so that by mid-1980's power consumption in agriculture was well above the one in the industrial sector [Dubash, 2007].

In my empirical specification, I use the night lights as the proxy of the quantity of electricity used for irrigation. According to smart power India (2019), electricity is mainly used for irrigation in rural areas. [6] For instance, the average electricity consumption for rural households is about 39.3 kWh/month and is used for general lighting, air circulation, household chores or entertainment. For rural enterprises the level of consumption is similar to the households' one (39.5 kWh/month) and is mainly used for general lighting, air circulation or productive use (62% of the total consumption). The consumption in the agricultural sector is almost the double (80 kWh/month) and is exclusively used for irrigation purposes. For this reason, I am confident in claiming that an increase in electrification may increase the comparative advantage of the agricultural sector compared to the non-agricultural sector. That is why also I restrict my sample to individuals/households that live in rural areas.

5. http://www.indiaenergy.gov.in/doc/Expert-view/Rural%20Electrification%20India.pdf

6. https://smartpowerindia.org/media/1230/report$_r$ural-electrification-in-india$_c$ustomer-behaviour-and-demand.pdf

7 DATA

I use several sources of data for this paper : the rural sample of the National Sample Survey (NSS) of India, PRIO-GRID, data from Allcott et al. [2016a], the Ministry of Agriculture and Farmers Welfare and the Central Ground Water Board (CGWB) of India.

National Sample Survey (NSS)

Variables related to the employment or unemployment of rural individuals/households come from the NSS. It is a repeated cross section database collected periodically by the Ministry of Statistics and Program Implementation of India and is representative at the state and district levels. The dataset contains informations on labor market activities of individuals, such as their principal and subsidiary activity status or if they work on a regular basis. One of the advantage in using NSS is that it allows to know the industry in which individuals work at the time of the survey, allowing to study labor reallocation between sectors between 2000 and 2004. Data are collected at the individual and household level. I aggregate them at the district level to form a panel database. The main variable of interest is the employment share in the agricultural and manufacturing sectors. Rounds of NSS that are used are the 55^{th} round collected in 2000 and the 61^{st} round collected in 2004. Table 3.1 shows descriptive statistics of the main variables, displayed separately for years 2000 and 2004. I notice a decrease of the agricultural labor share from 75% to 64%, i.e a fell by 11 percentage point. Much of this labor force seems to be absorbed by sectors other than the manufacturing. [7] Their combined labor share rose from 18% to 27% between 2000 and 2004, while it rose from 6% to

7. I refer to the National Industrial Classification (NIC) of 1998 to classify economic activities into broad sectors. The agricultural sector employs a large part of the labor force, followed by the manufacturing sector. I regroup the remaining sectors into the rubric "other sectors", given that their corresponding labor share is small. In total I identify 21 broad sectors.

8% for the manufacturing sector. This pattern in the data presents the premises of a movement of the labor from the agricultural sector towards non-agricultural sectors through time.

Agricultural outcomes

I obtain data on crop at the district-season-year level from the Ministry of Agriculture and Farmers Welfare. I aggregate agricultural production regardless of crop type or season to have total annual production by district in tons. I compute the variable output per worker by dividing the annual agricultural production by the number of people working in the agricultural sector. This variable allows to estimate the productivity of the agricultural labor. It is measured by the quantity of output produced per worker. In addition, this database contains information on the area cultivated in each district. It is used to construct the labor intensity variable by dividing the number of individuals working in the agricultural sector by the area cultivated. It is expressed in workers per hectare of land cultivated. Panel B of table 3.1 shows descriptive statistics of these variables. I mainly notice that there is not an important variation of these variables between 2000 and 2004.

PRIO-GRID

The PRIO-GRID database is a standardized spatial grid structure with global coverage at a resolution of 0.5 x 0.5 decimal degrees, that corresponds to a cell of roughly 55 x 55 kilometers at the Equator (3025 square kilometres area). [8] It contains spatially disaggregated data at the grid cell level. The main variables used from this dataset are the night-lights that is an average measure of night time light emission, area irrigated which measures the area equipped for irrigation within each cell (in hectares) and agricultural area that gives the percentage area of the cell covered by agricultural land. My main variables of interest are the night light intensity as a proxy of the quantity of electricity available within a district and the

8. `https://www.prio.org/Data/PRIO-GRID/`

area irrigated. Figure 2.1 shows that there is lot of variation between districts for the night lights and the area irrigated and visually, there appears to be a spatial correlation between these two variables. Figure 2.2 presents their Kernel density plots and shows their variation between 2000 and 2004. It mainly shows that there is not enough temporal variation for the area irrigated. For the night lights the t-test mean difference between 2000 and 2004 is significant with a t-statistics equal to 4.43 and it is not significant for area irrigated, with a t-statistic of -0.0796. For this reason, I use the night lights in my empirical specification as the proxy of the quantity of electricity used for irrigation, instead of the area of the district equipped with irrigation installation. In the following section, I show the relationship between these two variables with basic correlations in the data.

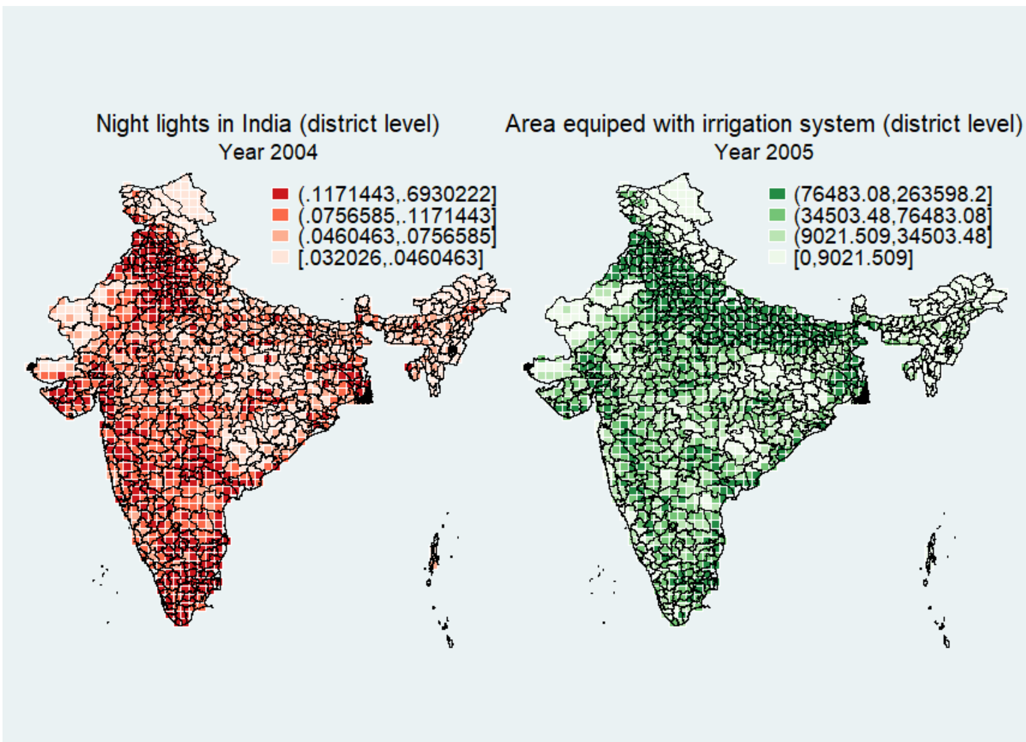
Table 2.1 presents also descriptive statistics of main covariates that I use in my specification. The average rainfall (in m^3) and temperature (in degree Celsius) are determinants of agricultural as well as non-agricultural production [Allcott et al., 2016a]. I also include socio demographic characteristics of districts such as the literacy and unemployment rate, the income per capita, the number of discriminated or powerless ethnic groups and the distance to the national capital city.

Finally, I use data from Allcott et al. [2016a]. I am interested on the instrumental variable they developed that is the predicted hydroelectricity generation as a share of predicted electricity demand at the state level.

Table 2.1:
Descriptive Statistics

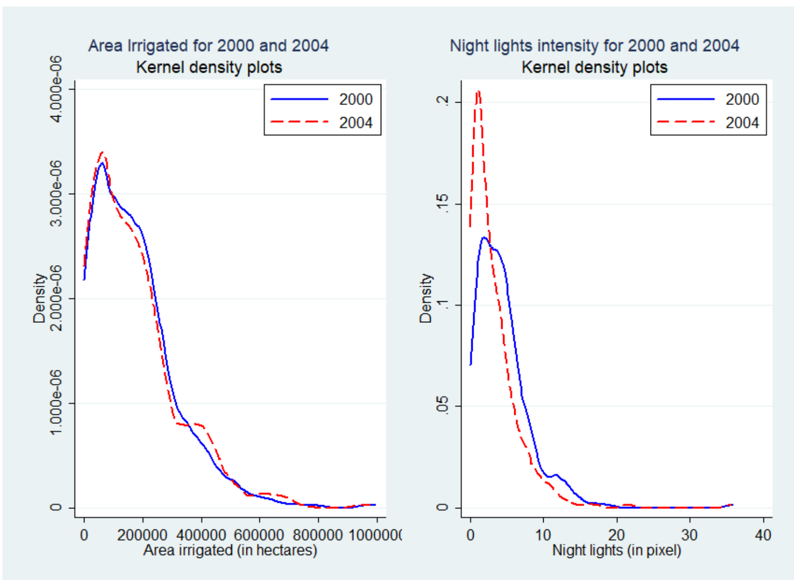
	2000		2004		t-test		Observations
	Mean (1)	SD (2)	Mean (3)	SD (4)	Diff (5)	t-Stat (6)	(2000 + 2004) (7)
Panel A : Employment Share							
Agricultural sector	0.75	0.14	0.64	0.11	0.11	12.23	718
Manufacturing sector	0.06	0.06	0.08	0.06	-0.02	-4.84	718
Other sectors	0.18	0.11	0.27	0.08	-0.09	-12.13	718
Panel B : Agricultural Outcomes							
log Output per worker (kg/Worker)	8.29	1.83	8.41	1.89	-0.11	-0.79	713
log Output (kg)	13.63	2.10	13.66	2.04	-0.03	-0.21	843
Labor intensity (Worker/ha)	2.19	1.58	2.17	4.25	0.03	0.1	716
log Agricultural Land (ha)	5.01	0.79	4.91	0.88	0.09	1.54	718
Panel C : Natural resources outcomes							
log Forest Cover (ha)	6.23	1.83	6.13	1.89	0.09	0.73	837
Depth to Water level (meters)	7.3	5.9	8.13	5.95	-0.83	-2.04	854
Panel D : Night Light Intensity & Hydroelectricity							
Night Lights	4.43	4.24	3.18	3.48	1.26	4.36	718
Hydroelectricity supply	0.24	0.25	0.18	0.25	0.06	3.31	718
Hydroelectricity * $E_{d,s,0}$	0.57	0.91	0.38	0.67	0.19	3.28	718
Panel E : Climate Data							
Rainfall (m^3)	3.04	2.04	3.25	2.36	-0.21	-1.25	718
Temperature (°C)	24.85	4.54	24.94	4.44	-0.09	-0.28	718
Panel F : Socio-Demographic Variables							
Literacy rate	0.43	0.15	0.51	0.14	-0.08	-6.87	718
Unemployment rate	0.31	0.12	0.32	0.1	-0.01	-1.73	718
Number of excluded groups	0.15	0.58	0.13	0.55	0.02	0.55	716
Distance to the National capital city	961.74	570.27	958	550.64	3.73	0.09	718

Notes : $E_{d,s,0}$ is considered as the initial level of electrification of districts (1992). In total, year 2000 includes 418 districts and year 2004 437 districts. Due to missing values for some variables, there is maximum 82% of districts in 2000 (345) and 85% in 2004 (375).



Note : Author's graphic from PRIO-GRID data

FIGURE 2.1 Spatial correlation between Night lights intensity and Area irrigated



Note : Author's graphic from PRIO-GRID data

FIGURE 2.2 Kernel density plots for Night lights intensity and Area irrigated

8 CORRELATION BETWEEN ELECTRICITY AND AREA IRRIGATED

In this section, I am interested on the relationship between the expansion of electricity and area irrigated. The objective of this correlation is to show some evidence that electricity is mainly used for irrigation purposes in rural India. The equation to be estimated is as follows :

$$log(AI_{gid,t}) = \delta_t + \delta_s + \beta_1 Electricity_{gid,t} + \beta_2 Temperature_{gid,t} + \epsilon_{gid,t} \tag{2.1}$$

where gid is the grid cell unit of observation in the PRIO-GRID database, t is time (1995, 2000 and 2005), δ_t and δ_s are respectively time and state fixed effects, $log(AI)$ is the logarithm of area irrigated in each cell and $Electricity$ represents the night light variable.

The estimated coefficient β_1 in Table 2.2 implies that a 1 unit increase in electricity corresponds to a 16.6% increase in area irrigated. This coefficient estimate, combined with Figure 2.1 show that areas with a greater availability of electricity experience a faster expansion of area irrigated. In the literature, some authors make the same observation, with Briscoe and Malik [2006] who argue that groundwater pumping use around 40% of total electricity consumption. State governments make lot of financial efforts to provide electric power to farmers as shown by Bhatia (2005) (Figure 20, p.11). He argues that the World Bank estimates that subsidies to farmers account for about Rs 240 billion a year, equivalent to 25 percent of India's fiscal deficit. In some states such as Gujarat, state governments put in place a feeder segregation program which objective is mainly to change the physical power supply infrastructure and to separate the non-agricultural and agricultural feeders [Chindarkara, Chena, and Satheb, 2017]. This program is associated with a management of electricity supply to farmers, including an improvement in the quality and reliability of electricity supply and an increase in the period of availability of electricity for agricultural activities. These evidences show that electricity supply for irrigation is of central importance in India.

In the next section, I present and implement an identification strategy that tries to establish causal effect of electricity expansion on labor reallocation between sectors and also on natural resources in India.

Table 2.2:
Correlation between electricity and area irrigated

	(1) log(area irrigated)
Electricity	0.166***
	(0.0209)
Temperature	0.143***
	(0.0252)
Constant	6.078***
	(0.629)
State F-E	Yes
Year F-E	Yes
Observations	3248

Notes : Standard errors in parentheses are clustered at the district level. Dependent and independent variables are considered over the years 1995, 2000 and 2005. The unit of observation is the grid cell. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

9 IDENTIFICATION STRATEGY

My empirical strategy seeks to study the impact of electrification on labor reallocation in rural areas in India and also on the forest cover and the water table. The following equation represents the baseline specification :

$$Y_{d,t} = \alpha_s + \gamma_t + \beta_1 E_{d,t} + \beta_2 X_{d,t} + \epsilon_{d,t} \tag{2.2}$$

Where $Y_{d,t}$ is an outcome variable measured in district d and year t (2000 and 2004). $E_{d,t}$ is the night light intensity considered as the level of electrification of districts. $X_{d,t}$ is a vector of control variables at the district level such as the average rainfall and temperature, the number of excluded groups defined as the number of powerless or discriminated ethnic groups in the district, the distance that separate the district to the national capital city, the distance in kilometer that separate the district to the border of the nearest land-contiguous neighboring country, the distance to the border of the nearest neighboring country, regardless of whether the nearest country is located across international waters and the shortest straight-line distance to international waters. α_s and γ_t are respectively state and year fixed-effects.[9] Finally, $\epsilon_{d,t}$ is the error term, clustered at the district level. My variable of interest is $E_{d,t}$. Its coefficient, β_1, captures the variation of the outcome variable (ΔY_d) following a variation in electrification (ΔE_d).

I am first interested on how electrification affects the labor and land factor in agriculture. This first step will allow to determine if electricity is a factor-saving or factor-biased technology [Bustos et al., 2016]. For example, a positive effect of electrification on land expansion would mean that farmers prefer to expand their area cultivated. It will be, therefore, interesting to see the consequences on natural resources [Angelsen and Kaimowitz, 2001]. If I consider the effect of electrification on the agricultural labor share, workers per unit of land or output per worker, the sign of the estimated coefficients would tell us whether there is an increase or a decrease of labor needed for agricultural activities. I will analyse then the repercussions on the labor share of non-agricultural sectors (manufacturing).

9. I use two rounds of the National Sample Survey with around 720 observations and 360 districts. A specification with district fixed-effects does not provide enough variation for identification. I, instead, use a state level fixed-effects (35 states). I include district level controls to take into account their specific characteristics.

An OLS estimation of equation (3.2) is with no doubt inconsistent with the endogeneity of the level of electrification of districts due to reverse causality or omitted variables. For instance, the level of electrification of some areas compared to others may be correlated to unobserved political or economic characteristics. I address the problem of endogeneity by using a Bartik instrument : the predicted hydroelectricity supply shifts developed by Allcott et al. [2016a], weighted by the "initial" electrification level of districts. The hydro supply shifts ($Hydro_{s,t}$), developed at the state level, represents the share of predicted electricity demand covered by state s predicted Hydroelectricity generation. [10] In their paper, authors use it to instrument electricity shortages. [11]

Hydroelectricity is the second source of energy in India behind coal. It is considered by the Indian government as more reliable and affordable that fossil fuels, because it is a renewable energy. [12] In addition, Indian coal power plants were offline 28% of the time between 1994 and 2009 [Chan, Cropper, and Malik, 2014]. Thus, the advantage using the hydroelectricity supply shifts as instrument is that it captures the quantity and quality of electricity supply. Applied to the context of this paper, I can say that a farmer will use electricity for irrigation purposes if and only if it is available in his area. Thus, the quantity

10.

$$Hydro_{s,t} = \frac{H_{s,t}}{\tilde{Q}_{s,t}} \tag{2.3}$$

Where $H_{s,t}$ is the predicted state-level hydroelectricity generation computed by using reservoir inflows and generation from "run-of-river" hydro plants that have no reservoirs to store water. And $\tilde{Q}_{s,t}$ represents the predicted electricity demand computed as follows :

$$\tilde{Q}_{s,t} = \sum_{r \neq s} Q_{r,t} \cdot \sum_{y=1992}^{2010} \frac{Q_{s,y}}{\sum_{r \neq s} Q_{r,t}} \tag{2.4}$$

By dividing predicted hydroelectricity generation by the predicted electricity demand we obtain the relative share of hydro generation across states.

11. Allcott et al. [2016a] are interested in estimating how variation in electricity shortages affects the performances of manufacturing plants in India. The possibility that economic growth may increase electricity demand and, therefore, electricity shortage, leads to an endogeneity problem coming from reverse causality. To address the problem, authors use the predicted supply shift of hydroelectricity generation as an instrument for shortages.

12. `https://www.energy.gov/eere/water/benefits-hydropower`

of electricity available, as well as the probability for a household to have access to electricity, can be related to the supply shifts from hydroelectricity generation in his home state. It is therefore possible to use the latter as an instrument for electricity in agricultural activities, with the assumption that the instrumental variable may influence agricultural outcomes only through the use of electricity for irrigation.

A threat to this assumption is that there is the possibility for a hydroelectric central to store water in reservoir for future years if demand is low in the present year. Actual hydroelectricity generation would, in this case, be related to the demand of electricity. That is the reason why Allcott et al. [2016a] use the predicted hydroelectricity supply and the predicted electricity demand to compute the share of the electricity demand satisfied by hydro generation. [13]

Since I intend to do a district-level analysis, I transform the state level instrument of Allcott et al. [2016a] to a district level one.

$$Hydro_{d,s,t} = Hydro_{s,t} * E_{d,s,0} \tag{2.5}$$

Where $E_{d,s,0}$ represents the "initial" level of electrification of a district d in state s. [14] Given that data on night lights are available from 1992 to 2013, I consider light intensity in 1992 as the initial level of electrification of districts.

Equation (2.6) represents the first stage specification :

13. For more information about how the state-level instrument has been computed, see Allcott et al. [2016a].

14. The necessary condition for doing this transformation is to ensure that the new instrumental variable at the district level ($Hydro_{d,t}$) is still valid ; that is to say $E(Hydro_{d,t} * \epsilon_{d,t}) = E(Hydro_{s,t} * E_{d,s,0} * \epsilon_{d,t}) = 0$. Taking the limit of this expression, we can have : $\lim_{S,D\to\infty} \frac{1}{D*S} \sum_D \sum_S (Hydro_{s,t} * E_{d,s,0} * \epsilon_{d,t}) = 0$. $\lim_{S,D\to\infty} \frac{1}{D*S} \sum_S Hydro_{s,t} \sum_D (E_{d,s,0} * \epsilon_{d,t}) = 0$. And this condition is verified if and only if $\lim_{D\to\infty} \frac{1}{D} \sum_D (E_{d,s,0} * \epsilon_{d,t}) = 0$ or $E(E_{d,s,0} * \epsilon_{d,t}) = 0$. This condition means that shocks related to the agriculture and manufacturing sector in 2000 and 2004 are not correlated to districts night light intensity in 1992.

$$E_{d,t} = \alpha_s + \gamma_t + \phi_1 Hydro_{s,t} * E_{d,s,0} + \phi_2 Hydro_{s,t} + \phi_3 X_{d,t} + \mu_{d,t} \tag{2.6}$$

With $Hydro_{s,t} * E_{d,s,0}$ instrumenting $E_{d,t}$.

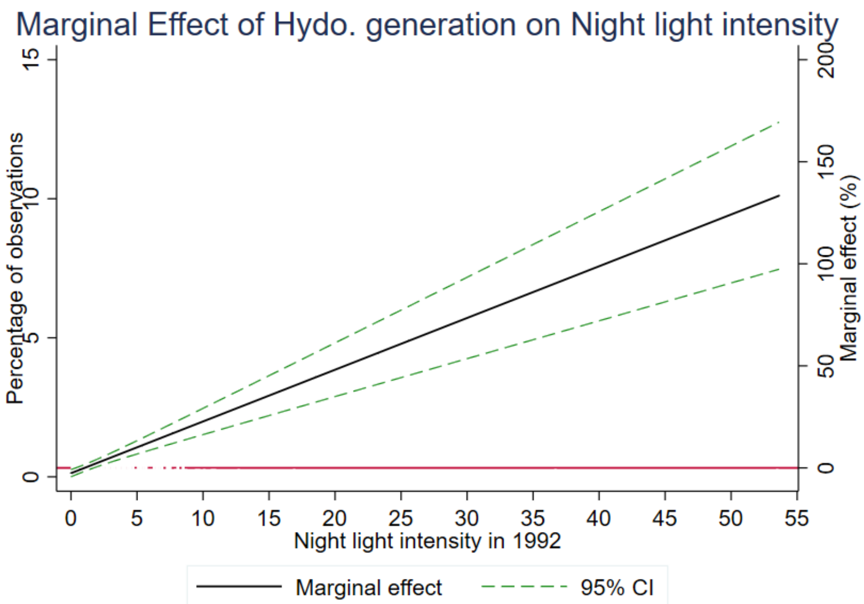
Table 2.3 presents the results of the first stage specification. It evaluates the hydro instrument, controlling for rainfall, temperature, socio-economic variables as well as state and year fixed-effects. I present 3 different specifications with changes in the set of control variables. Column (1) displays the regression results of electricity on the instruments without control variables, column (2) includes the average temperature and rainfall that are determinant for hydroelectricity production, while column (3) includes all the set of controls. We can see that coefficients of interest are stable accross specifications. Column (3) is the most conservative one and it shows that a one unit shift in the Hydroelectricity supply increases night lights by 2.69 units for districts with a higher level of initial electrification. Min and Gaba [2014] find that a 1 unit increase in night lights is equivalent to 60 public street-lights. Applying this numbers to this study I can say that the coefficient estimate of 2.69 units correspond to around 161 street-lights. This result suggest that the district level instrument has the expected positive sign on the night light intensity. The instrument at the state level has to be interpreted as a marginal effect with the coefficient on the interaction term. Figure 2.3 shows that an hydro supply shift has a positive effect on the night light intensity of districts that have a high initial level of electrification and it is negative for districts with a low initial level. As explained by Gupta and Pelli [2021], a high initial level of electrification allows to use more effectively the positive shifts, while no initial electrification would render them useless. Given that I cluster the error term at the district level, I present the Kleibergen-Paap F-Statistics. In column (3) it is equal to 26.83, suggesting that my instruments are relevant. The P-value of the Hansen J statistic is equal to 0.45 and it validates the over-identifying restriction.

TABLE 2.3:
First stage with Hydro as Instrument

	(1)	(2) Electricity	(3)
$Hydro_{s,t} * E_{d,s,0}$ ♠	2.76***	2.7***	2.69***
	(0.38)	(0.38)	(0.35)
$Hydro_{s,t}$ ♣	-2.98***	-2.52**	-2.43***
	(0.83)	(0.92)	(0.9)
Temperature		0.11*	0.113**
		(0.05)	(0.05)
Rainfall		-0.07	-0.07
		(0.06)	(0.07)
Distance 1			0.0006
			(0.001)
Distance 2			0.0005
			(0.92)
Distance 3			-0.003*
			(0.82)
Number of excluded groups			-0.131
			(0.18)
Distance to the national capital city			-.003***
			(0.0007)
State F-E	Yes	Yes	Yes
Year F-E	Yes	Yes	Yes
Observations	716	714	712
Kleibergen-Paap F-Stat	26.01	24.87	26.83
Hansen J stat. (P-value)	0.08	0.12	0.455

Notes : Standard errors are clustered at the district level. The dependant variable is the night lights intensity as a proxy of the level of electrification of districts. F-statistic is for the heteroskedasticity and cluster-robust Kleibergen-Paap weak instrument test. ♠ is the district level instrument for the night lights intensity $Hydro_{d,t}$. ♣ is Allcott et al. [2016a] instrument at the state level. *Distance 1* represents the distance in kilometer that separate the district to the border of the nearest land-contiguous neighboring country. *Distance 2* is the distance to the border of the nearest neighboring country, regardless of whether the nearest country is located across international waters. Finally, *Distance 3* is the shortest straight-line distance to international waters.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



Note : Author's graphic

FIGURE 2.3 First stage marginal effect for the hydro supply shift on night lights

10 EMPIRICAL RESULTS

First, I present the effect of electricity for irrigation on agricultural outcomes. I show evidence on land and labor in order to know if electricity is a factor-biased or factor-saving technology. Second, I look at the repercussions on the manufacturing sector, in terms of labor reallocation. I, then, look at the effect on forest cover and depth of the water table. I also perform an analysis of labor reallocation between the agricultural and non-agricultural sectors at the household level in order to validate the district level analysis.

10.1 Effect on agricultural outcomes

Table 2.4 reports the results for agricultural outcomes. Column (1) shows how land is affected, while column (2) and (3) show evidences on labor. Column (4) and (5) show how electricity improves agricultural output. Column (1) suggests that electricity use for irrigation has a statistically significant effect on land expansion. The point estimate is significant at 10% level. All else equal, a unit increase in night lights predicts a decrease of the area cultivated between 2000 and 2004 by 4.7%. This finding supports the hypothesis that electricity use in agriculture in India is land-saving.

I then analyse the effect of electricity on the labor factor. In Column (2) of Table 2.4 I run the regression of equation (2.2) with the employment share of the agricultural sector as the outcome variable. The estimate obtained suggest that the employment share decrease following an increase of night lights. The point estimate implies that a unit increase in night lights is combined with a decrease of the employment share in the agricultural sector by 0.022. This is equivalent to a decrease by 2.2%. Column (3) indicates that the number of agricultural workers used per unit of land (labor intensity) increases by 0.117 individuals per hectare. This increase of the labor intensity may be explained by the fact that small-scale farmers mainly employ family members for farming activities. A reallocation of part of this family labor towards other sectors may imply an adjustment of the amount of labor needed per hectare cultivated. Farmers may, therefore, reduce there area cultivated in order to have an optimal production. Despite the decrease of the quantity of factors of production (labor and land) used for agriculture, column (4) and (5) indicate that areas with fastest development of electrification experience a greater increase in agricultural output. A unit increase in night lights is synonym of an increase of agricultural output by 13% and the output per worker by 10.8%.

Taken all together, the results presented in Table 2.4 suggest that electrification has a considerable impact in the agricultural production process with labor and land factors that are affected. Areas where the use of electricity for irrigation is greater, experience a greater increase of agricultural output, a reduction in the area cultivated and in the agricultural labor share. These findings are consistent with the hypothesis that suggest that electricity for irrigation is labor and land saving. They also shows the benefits of electrification as it increases farmers' production with less resources required. Given that agriculture has priority in agricultural household consumption, the increase of agricultural output may allow them to have a sufficient amount of food and to release labor for non-agricultural activities.

In the remainder of the analysis, I show whether this outflow of labor from the agricultural sector is redeployed in other sectors of the economy. With a population of almost 1.4 billion of individuals, the labor reallocation between sectors in India may involve millions of individuals, generating a structural transformation of the economy. I also look at the consequences on natural resources (forest cover and water table) as a consequence of the land-saving technology.

Table 2.4:
Agricultural outcomes

	Log Land cultivated		Empl. share		Labor intensity		log Output		log output/worker	
	IV (1)	OLS (2)	IV (3)	OLS (4)	IV (5)	OLS (6)	IV (7)	OLS (8)	IV (9)	OLS (10)
Electricity	-0.047**	-0.020*	-0.022***	-0.008***	0.117***	0.045***	0.130***	0.034*	0.108***	0.019
	(0.022)	(0.012)	(0.004)	(0.001)	(0.036)	(0.012)	(0.041)	(0.019)	(0.034)	(0.016)
Temperature	0.033**	0.029*	0.003	0.001	-0.076**	-0.064**	0.018	0.034	0.003	0.018
	(0.016)	(0.016)	(0.003)	(0.002)	(0.032)	(0.031)	(0.022)	(0.021)	(0.021)	(0.019)
Rainfall	-0.017	-0.017	-0.003	-0.002	0.122**	0.118**	-0.029	-0.034	-0.017	-0.022
	(0.025)	(0.025)	(0.004)	(0.004)	(0.047)	(0.048)	(0.034)	(0.035)	(0.028)	(0.028)
Distance 1	-0.001***	-0.001***	0.000	0.000	0.001	0.001	-0.000	0.000	0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Distance 2	-0.000	-0.000	-0.000	0.000	0.001	0.001	0.002**	0.001**	0.001*	0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Distance 3	0.001**	0.001***	-0.000	0.000	-0.002***	-0.002***	-0.003***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of excluded groups	0.117	0.128	0.008	0.011	-0.087	-0.103	0.135	0.112	-0.013	-0.035
	(0.086)	(0.084)	(0.011)	(0.011)	(0.105)	(0.104)	(0.094)	(0.100)	(0.101)	(0.100)
Distance to the national capital city	0.000*	0.000**	-0.000***	-0.000*	0.001**	0.000	-0.001***	-0.001***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
State F-E	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F-E	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	710	710	712	712	709	709	709	709	708	708

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the district level. Dependant variables are measured at the district level. All regressions include state and year fixed-effects. *Electricity* is the night lights intensity as a proxy of the level of electrification of districts. *Distance 1* represents the distance in kilometer that separate the district to the border of the nearest land-contiguous neighboring country. *Distance 2* is the distance to the border of the nearest neighboring country, regardless of whether the nearest country is located across international waters. Finally, *Distance 3* is the shortest straight-line distance to international waters.

10.2 Effect on non-agricultural sectors

I study the effect of electrification on the labor share of the manufacturing sector and all other sectors combined together. Table 2.5 presents the results. In column (1), I notice an increasing labor share in the manufacturing sector. The point estimate suggests that a unit increase in night lights is accompanied with an increase of the manufacturing employment share in rural areas by 0.8%. Regarding other non-agricultural sectors, I notice an increase of their combined labor share by 1.5% following a unit increase in night lights (Column 3).

Table 2.5:
Non-agricultural sectors' employment share

	Manufacturing		Other sectors	
	IV (1)	OLS (2)	IV (3)	OLS (4)
Electricity	0.008***	0.003***	0.015***	0.005***
	(0.002)	(0.001)	(0.003)	(0.001)
Temperature	-0.002	-0.001	-0.002	-0.000
	(0.001)	(0.001)	(0.002)	(0.002)
Rainfall	0.003	0.002	0.000	-0.000
	(0.002)	(0.002)	(0.003)	(0.003)
Distance 1	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Distance 2	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Distance 3	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Number of excluded groups	0.003	0.002	-0.011	-0.013
	(0.004)	(0.004)	(0.009)	(0.009)
Distance to the national capital city	0.000***	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
State F-E	Yes	Yes	Yes	Yes
Year F-E	Yes	Yes	Yes	Yes
Observations	712	712	712	712

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the district level. Dependant variables are measured at the district level. All regressions include state and year fixed-effects. *Electricity* is the night lights intensity as a proxy of the level of electrification of districts. *Other sectors* regroups all sectors other that agricultural and manufacturing sectors. *Distance 1* represents the distance in kilometer that separate the district to the border of the nearest land-contiguous neighboring country. *Distance 2* is the distance to the border of the nearest neighboring country, regardless of whether the nearest country is located across international waters. Finally, *Distance 3* is the shortest straight-line distance to international waters.

10.3 Diversification of household activities

In this section, I perform a repeated cross-section regression at the household level in order to see if there is a diversification of household activities following an increase of electrification. As said above, a farmer using a traditional method for watering crops may require all family members for farming activities. Following the use of electricity, less labor will be needed with the electric pump that do the heavy duties. In this case, one can expect a reallocation of the surplus of family members towards other type of activities in other sectors and a diversification of the household income sources. This analysis at the household level will allow to confirm the results found with the district level analysis. I restrict my sample to rural households where the head is working in the agricultural sector and create a variable that count the number of household members not working in the agricultural sector. I then run the following specification :

$$Y_{h,d,t} = \alpha_s + \gamma_t + \beta_1 E_{d,t} + \beta_2 X_{d,t} + \beta_3 H_{h,d,t} + \epsilon_{h,d,t} \tag{2.7}$$

Where $Y_{h,d,t}$ is the number of household members not working in the agricultural sector as the household head, α_s and γ_t are respectively state and year fixed-effects, $E_{d,t}$ is the electricity variable at the district level. $X_{d,t}$ represents the average temperature and rainfall at the district level. $H_{h,d,t}$ is a vector of control variables at the household level that includes the household head gender and education level.

The results shown in Table 2.6 are consistent with the district level analysis. They mainly suggest that in rural households where the head operates in the agricultural sector, there is a diversification of household activities. The point estimate in column (1) indicates that there is an increase of non-agricultural workers at the household level by 0.011 individuals. It is equivalent to an increase of 6.11% compared to the mean of household's non-agricultural workers. Obviously I note the opposite effect for agricultural workers by around the same magnitude, although the point estimate is not significant.

Taken together, conclusions of the household and district level analysis recommend that households diversify their income sources following the use of electricity for irrigation.

Table 2.6:
Effect of electricity on household activities diversification

	# Non-agri. workers		# agri. workers	
	(1)	(2)	(3)	(4)
Electricity	0.011***	0.004***	-0.012	-0.002
	(0.003)	(0.001)	(0.009)	(0.004)
Temperature	0.000	0.002	0.001	-0.001
	(0.002)	(0.002)	(0.007)	(0.006)
Rainfall	0.004	0.005	-0.003	-0.004
	(0.004)	(0.004)	(0.010)	(0.010)
log(Land Cultivated)	0.019***	0.018***	0.137***	0.138***
	(0.001)	(0.001)	(0.005)	(0.005)
Head-Gender	-0.014	-0.014	-0.449***	-0.449***
	(0.010)	(0.010)	(0.025)	(0.025)
Head-Educ	Yes	Yes	Yes	Yes
State F-E	Yes	Yes	Yes	Yes
Year F-E	Yes	Yes	Yes	Yes
Observations	58392	58392	58392	58392
Mean of dep. var.	0.18		1.07	
	(0.52)		(1.22)	

Notes : The unit of observation is the household. Standard errors in parentheses are clustered at the district level. Column 1 and 2 consider as dependant variable household members that do not work in the agricultural sector. Column 3 and 4 consider household members that operate in the agricultural sector. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

10.4 Effect on natural resources outcomes

Between 2000 and 2004, results show that there is a decline of area cultivated. As a result, I expect a positive or neutral relation between the use of electricity for irrigation and the forest cover. Results presented in Table 2.7 show the opposite with a decrease of forest cover following an increase of electrification. The estimated coefficient in column (1) indicates that a 11.4% decrease of forest cover can be associated to a unit increase in night lights. The effect that I capture here is probably due to the fact that the main source of electricity generation in India is coal and its effects on natural resources such as forest cover are negative and may lead to deforestation or forest degradation. In fact Bebbington, Bebbington, Sauls, Rogan, Agrawal, Gamboa, Imhof, Johnson, Rosa, Royo, et al. [2018] explain that coal mining has a direct effect on forest cover through greenhouse gas emission. It also has an indirect effect through infrastructure development, in-migration, new human settlements, and other economic activities.

Column (3) of Table 2.7 investigates the effect of electrification on the water table. The estimate is statistically not significant. This finding is opposed to the literature on groundwater pumping in India [Badiani and Jessoe, 2013, Fishman et al., 2015, Fishman, Siegfried, Raj, Modi, and Lall, 2011]. Using electricity for irrigation has the potential to reduce water over extraction by adopting technologies like drip or sprinkler for watering plants. By considering all Indian states, I find no negative impact of electricity on groundwater extraction. And the magnitude of the coefficient is small because it suggests that a one unit increase in electrification implies a 0.024 meters (2.4 centimetres) decrease of the water table.

Table 2.7:
Effect of electricity on natural resources

	log(Forest cover)		Depth to Water level	
	IV (1)	OLS (2)	IV (3)	OLS (4)
Electricity	-0.114***	-0.091***	0.024	0.168
	(0.036)	(0.018)	(0.113)	(0.141)
Temperature	-0.057**	-0.061**	0.053	0.027
	(0.027)	(0.026)	(0.072)	(0.072)
Rainfall	0.136***	0.136***	-0.521***	-0.515***
	(0.031)	(0.031)	(0.105)	(0.107)
Distance 1	0.001**	0.001**	-0.001	-0.002
	(0.001)	(0.001)	(0.002)	(0.002)
Distance 2	0.000	0.001	-0.009***	-0.009***
	(0.001)	(0.001)	(0.002)	(0.002)
Distance 3	0.001	0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.003)	(0.003)
Number of excluded groups	0.152*	0.157**	-0.025	0.053
	(0.079)	(0.079)	(0.318)	(0.347)
Distance to the national capital city	-0.000	0.000	-0.004**	-0.003*
	(0.001)	(0.001)	(0.002)	(0.002)
State F-E	Yes	Yes	Yes	Yes
Year F-E	Yes	Yes	Yes	Yes
Observations	704	704	683	683
Mean of dep. var.	6.19		7.72	
	(1.95)		(5.94)	

Notes : The unit of observation is the district. Standard errors in parentheses are clustered at the district level. All regressions include state and year fixed-effects. *Electricity* is the night lights intensity as a proxy of the level of electrification of districts. *Distance 1* represents the distance in kilometer that separate the district to the border of the nearest land-contiguous neighboring country. *Distance 2* is the distance to the border of the nearest neighboring country, regardless of whether the nearest country is located across international waters. Finally, *Distance 3* is the shortest straight-line distance to international waters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The instrumental variable approach adopted in this project allows me to deal with the problem of endogeneity. However, the other problem that I can face is how to highlight the channel through which the use of electricity can have an effect either on the structural transformation of the Indian economy or on the exploitation of its natural resources. The aim of this paper is to estimate these effects only through a shock of agricultural productivity and labor reallocation. As pointed out by Colmer [2021] it can be challenging to isolate the effect of electricity use on manufacturing and natural resources outcomes only driven by a shock of productivity in the agricultural sector and a reallocation of the labor. In fact, this hypothesis supposes that the effect estimated does not take into account other potential channels through which the link can be established. [15]

To deal with this problem a solution is to adopt the strategy used by Colmer [2021] that considers the variation of the Indian labor market rigidity across states that amended the Industrial Dispute Act in a pro-worker direction and states that did it in a pro-employer direction. As explained by Besley and Burgess [2004b], this act aimed at protecting workers in the organized sector by increasing for example firms firing cost. Besley and Burgess [2004b] show that states that amended the Industrial Dispute Act in a pro-worker direction experienced a slowdown in economic activities (lowered output, input, investment and productivity) while those one which enacted pro-employer labor regulation had a better business climate and experienced a good growth performance. By giving more bargaining power to employees, pro-worker states have made entrepreneurs afraid of investing and employing workers. Thus,

15. Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi [2012] showed that with interconnected sectors, significant aggregate fluctuations may originate from a microeconomic productivity shock in one sector. In this case, one can suppose that electricity use may affect agricultural yields which in turn affects manufacturing firms that use agricultural goods as input. Santangelo [2016] describes another channel through which agricultural productivity may affect the manufacturing sector with a negative agricultural shock that may affect farmers wages that in turn may affect local demand of manufactured goods.

by using this variation in labor regulation across states, I can expect a limited effect of labor reallocation on structural transformation for pro-worker states compared to pro-employer ones. Equation (8) is the specification that I use.

$$Y_{d,t} = \alpha_s + \gamma_t + \beta_0 E_{d,t} + \beta_1 E_{d,t} * LR_s + \beta_2 X_{d,t} + \epsilon_{d,t} \tag{2.8}$$

The variable Labor Regulation (LR) takes the value 1 for pro-employer, 2 for neutral and 3 for pro-worker states. So there are 6 states that are considered pro-employer, 4 as neutral and 14 as pro-workers. [16].

However, it is important to consider the potential endogeneity resulting from the use of labor regulation to highlight the transmission channel. Indeed, as explained by Besley and Burgess [2004b], economic performances in the manufacturing sector of certain states may explain the adoption of the Act in a pro-worker or pro-employer direction. [17] This then raises a problem of reverse causality that risks skewing the results obtained. That is the reason why I do not interpret the results casually. I just use them to describe the pattern of labor reallocation for states according to their Industrial Dispute Act status.

Table 2.8 shows the results of the regression of equation (2.8). I notice a direct negative effect of electricity on the agricultural employment share as in Table 2.4. This negative direct effect correspond to the effect that we may have for pro-employer states. [18] It means that in this kind of states where the bargaining power of employers is greater than

16. Rajasthan, Madhya Pradesh, Andhra Pradesh, Karnataka, Kerala and Tamil Nadu are considered as pro-employer states. West Bengal, Orissa, Gujarat and Maharastra are neutral states. And Himachal Pradesh, Punjab, Chandigarh, Haryana, Delhi, Uttar Pradesh, Bihar, Sikkim, Arunachal Pradesh, Manipur, Mizoram, Tripura, Meghalaya and Assam are pro-worker states. This variable comes from Edmonds et al. [2010a]

17. In fact, good economic performances can push workers to claim more protection and benefits such as rights to a healthy work environment, the opportunity to participate in the capital of the enterprise or...

18. Pro-employer states are considered as the missing category in the regression of equation (2.8).

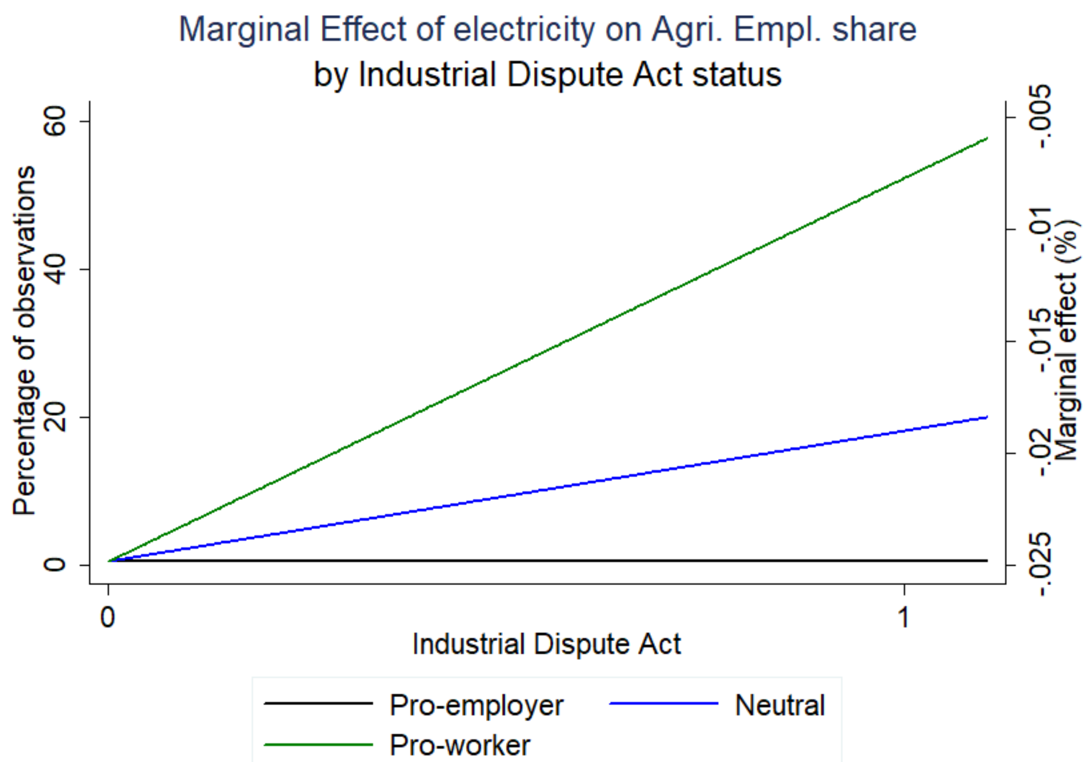
the one of workers in the manufacturing sector, we should have a decreasing trend for the agricultural labor share (column 1) and an increasing one for the manufacturing labor share (column 2).

Table 2.8:
Effect of electricity on labor reallocation subject to the Industrial Dispute Act

	(1) Agri. Empl. share	(2) Manu. Empl. share
Electricity	-0.018***	0.007***
	(0.005)	(0.002)
Electricity * $LR_{neutral}$	0.002	-0.007
	(0.008)	(0.006)
Electricity * $LR_{pro-worker}$	-0.012*	0.005
	(0.007)	(0.003)
Temperature	0.001	-0.000
	(0.003)	(0.001)
Rainfall	-0.001	0.002
	(0.005)	(0.002)
Distance 1	0.000	-0.000
	(0.000)	(0.000)
Distance 2	-0.000	0.000
	(0.000)	(0.000)
Distance 3	0.000	0.000
	(0.000)	(0.000)
Number of excluded groups	0.010	0.006*
	(0.011)	(0.003)
Distance to the national capital city	-0.000***	0.000***
	(0.000)	(0.000)
State F-E	Yes	Yes
Year F-E	Yes	Yes
Observations	649	649

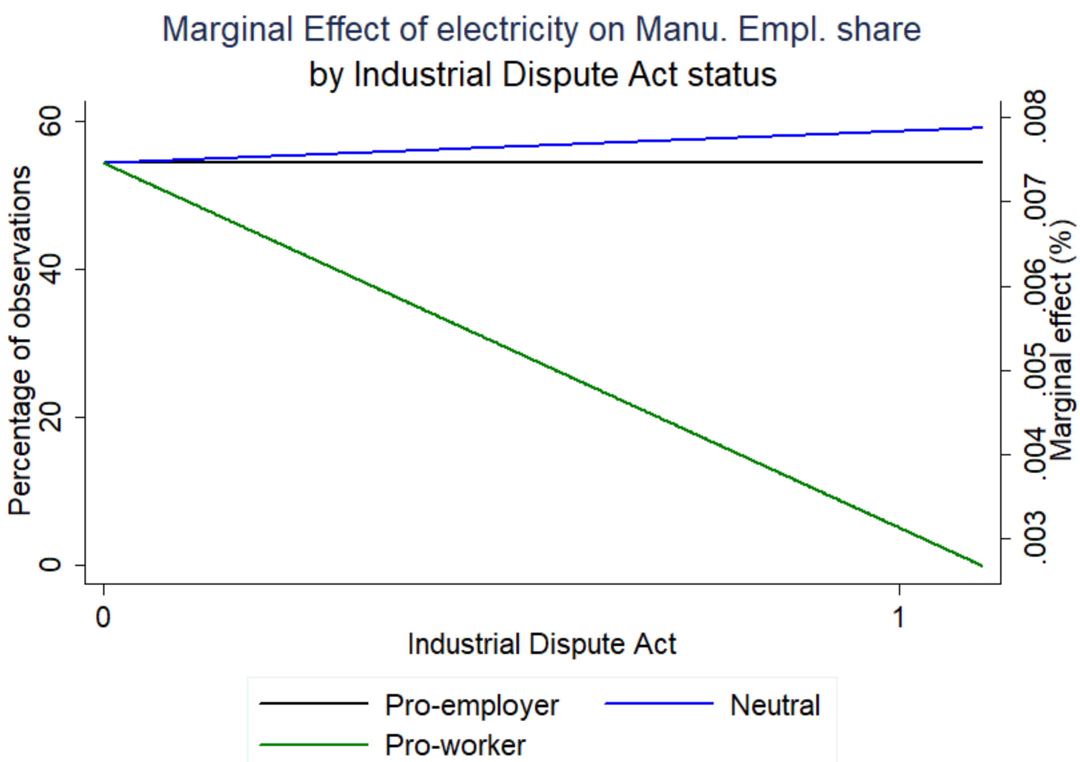
Notes : Standard errors in parentheses and are clustered at the district level. The pro-employer variable is the missing category. All regressions include state and year fixed-effects. *Electricity* is the night lights intensity as a proxy of the level of electrification of districts.*Distance 1* represents the distance in kilometer that separate the district to the border of the nearest land-contiguous neighboring country. *Distance 2* is the distance to the border of the nearest neighboring country, regardless of whether the nearest country is located across international waters. Finally, *Distance 3* is the shortest straight-line distance to international waters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 2.4 and 2.5 show what we should have in pro-worker and neutral states compared to pro-employer ones. The green line show that labor reallocation is not in the expected way for pro-worker states. In fact, the weak bargaining power of employers in these kind of states makes them reluctant to employ the labor force that should be released by the agricultural sector.



Note : Author's graphic

FIGURE 2.4 Marginal effect of electricity on the agricultural labor share



Note : Author's graphic

Figure 2.5 Marginal effect of electricity on the manufacturing labor share

12 CONCLUSION

Since the green revolution, the Indian government is doing lot of efforts to provide electricity to small-scale farmers for their irrigation needs. This program is a good example of how the adoption of a new agricultural technology can affect the development of an economy and also the exploitation of natural resources. In this article, I show empirical evidence on the effects of electricity use for irrigation on the process of structural transformation of the Indian economy and also its repercussions on the forest cover and the water table depletion.

Electricity allows farmers to increase their agricultural yields with less labor and land required in the agricultural production process. The extra workforce in the agricultural sector is then reallocated toward other sectors of the economy, such as the manufacturing sector. This results allow me to say that electricity use for irrigation favours both the development of the agricultural and manufacturing sector in India. My findings are in accordance with the literature on the effects of agricultural productivity on industrialization.

The nature of the link between agricultural productivity and natural resources is less obvious. By allowing the increase of agricultural production with less resources needed, one can expect less pressure on natural resources following the adoption of a new agricultural technology. But when we take into account the hypothesis of rationality of farmers, one can expect an overuse of natural resources. My results show no effect of electricity use for irrigation on water table depletion in India and the effect found on forest cover is probably due to electricity generation by Coal-fired power station.

In view of the results I have obtained, access to electricity for irrigation should be imperative for all Indian farmers because of its potential to favour the development of the Indian economy. But for this development to be sustainable, it is necessary to use renewable energy for powering the agricultural sector.

13 APPENDIX : NSS DATABASE CONSTRUCTION

I use several rounds on Employment and unemployment of the National Sample Survey (NSS) of India : NSS 43 (1987-1988), NSS 50 (1993-1994), NSS 55 (1999-2000), NSS 61 (2004-2005), NSS 64 (2007-2008), NSS 66 (2009-2010) and NSS 68 (2011-2012). NSS data are collected at the individual and household level and are a repeated cross section data. For the purpose of my work, I aggregate data at the district level to have a panel database. For that, I use many steps.

Extraction

I convert NSS data form ASCII to Stata format using the program developed by "Olivier Dupriez, World Bank Data Group (Nov. 2011)". NSS rounds are extracted separately. Data are grouped into text files, depending on whether they are at the individual or household level. For example, for NSS 66, files LV661001, LV661002, LV661003, group information at the hh level, while LV661004, LV661005, LV661006, LV661007, LV661008, LV661009, LV661010, LV661011 contain information at the individual level. Information contained in each file is extracted and put into in Stata format. I then merge them using two steps :

- I first merge database containing individual information by using key variables like the household and individual identifiers.
- The database obtained is then merged to database containing household information by using the household identifier.

It should be noted that there are duplicated observations for one or many members of a household. This is the case for family members that have two (work) activities during the week. One observation that specify the first activity, the number of days worked in that activity during the week. The second observation gives information about the second activity. All other variables remain the same (age, relation to the household's head, gender, etc.). For the first paper of my thesis I did not need these kinds of information. For this reason, I eliminate these duplicates when I aggregate at the district level in order not to consider the same individual many times. For my two other papers I do not eliminate these duplicates because I do the analysis at the individual level.

State and district code harmonization for rounds 55 and 43

There is a state and district code harmonization for NSS rounds 68, 66, 64 and 61. Codes are different for rounds 55 and 43. This is why I change districts and state codes for these 2 rounds. These dofiles are directly incorporated in the dofile extraction of NSS 43 and NSS 55 respectively.

District harmonization across NSS rounds

Administrative division in India is constantly changing with new districts being created using part of the territory of one or many other districts. That is why NSS rounds do not have the same number of districts and their geographic area is also not the same. I had to do a harmonization, by merging districts that were split to their districts of origin, to form a panel database. For example, NSS 43 contains around 441 districts and NSS 68 around 626 districts. I do the harmonization according to the administrative division of NSS 43, since it is the most distant year of NSS rounds. In fact, it is easier for me to merge districts that were split from their district of origin in more recent rounds (61, 64, 66, and 68) than creating ”new” districts that were carve out in oldest rounds (43, 50 and 55). I document a word file that explain when and from what (old) district a (new) district has been created. Information from the word document come mainly from Wikipedia. I also document an excel file that shows differences between NSS rounds. I create a code that performs the district harmonization for rounds 55, 61, 64, 66 and 68. There is no district code in NSS 50. For this reason, I do not use it in the harmonization. To do that, I just replace the district code of the (new) district by the district code of its district of origin. Here is an example : replace district = 16 if district == 19 & state == 3 : Banarla (district 19) is merged to Sangrur (district 16) in the state of Punjab (state 3). There are some specific cases where I was not able to identify the district of origin of a (new) district because it was carved out of many other districts. I did not know to what district of origin to merge it to. For these cases I did nothing. This is an explanation of the unbalanced panel. An example : S.A.S. nagar (Mohali) : It was formed in April 2006. This district was carved out of areas falling in Ropar (Rupnagar) and Patiala District.

Data aggregation

For the purpose of my work, information that seem important to aggregate at the district level are the proportion of individual working in each sector, the proportion of individuals working in each type of activity (principal and subsidiary activity), proportion of individuals seeking for work, proportion of individuals seeking for additional and alternative work and the average salary per industry. These variables can be obtained by applying weights or not. I use the variable "weekly activity nic" to compute the proportion of individuals working in each sector per district. In fact, the National Industry Classification (NIC) is used in the NSS to classify individuals by industry. I proceed as follows :

- extract the 2 first digit of weekly activity nic variable ;
- create a categorical variable with each category representing an industry based on the on the classification of NIC ;
- each NSS use a different NIC. So do a harmonization of the classification according to NSS 68 that use the most recent NIC (2008) ;
- each category (industry) is transformed as a dichotomic variable, as well as other categorical variables like the principal and subsidiary activity status of individuals, education level of individuals.

— when I use the sample weights, I first aggregate at the hh level. It allows me to apply household weights. And I then aggregate at the district level.

CHAPITRE 3

TRADE POLICY, AGRICULTURAL PRODUCTIVITY AND STRUCTURAL TRANSFORMATION : AN EMPIRICAL STUDY

1 AVANT-PROPOS

Je suis le seul auteur sur cet article. Le projet de recherche, le traitement de la base de données, la production des résultats ainsi que la rédaction sont le fruit de mon travail, sous la supervision de Pr. Pelli.

2 RÉSUMÉ

Dans cet article, j'étudie l'effet de la politique commerciale sur la relation entre la productivité agricole et la transformation structurelle en Inde rurale. La littérature comparent principalement la trajectoire de transformation structurelle de pays ayant des contextes socio-économiques différents et à des périodes différentes. Cet article considère les districts indiens comme de petites économies ouvertes et compare leur trajectoire de transformation structurelle suite à la réforme commerciale indienne de 1991 et à l'expansion de l'utilisation de l'électricité pour l'irrigation en Inde rurale. J'utilise l'interaction entre le tarif national ad-valorem et la part de consommation au niveau du district de chaque bien qui compose le panier de consommation des ménages ruraux comme mesure du tarif au niveau du district. La stratégie d'identification tient compte de l'endogénéité potentielle de la mesure du tarif et de l'électricité pour l'irrigation qui sont respectivement instrumentés par des variables d'accessibilité géographique d'un district donné et le le changement de l'offre hydroélectrique au niveau du district. Les résultats suggèrent qu'une baisse du niveau des tarifs, combinée à l'expansion de l'électricité, a un effet statistiquement significatif sur la réallocation du travail entre les secteurs. La part de la main-d'oeuvre agricole diminue dans les districts qui ont connu une baisse plus importante du niveau des tarifs et une augmentation de la part de la main-d'oeuvre dans les autres secteurs de l'économie. Conformément à la littérature, le secteur des services (commerce de détail, éducation et services publics) semble absorber la majeure partie de la main-d'oeuvre du secteur agricole.

3 ABSTRACT

In this paper, I study the effect of trade policy on the relationship between agricultural productivity and structural transformation in rural India. Papers in the literature mainly compare the structural transformation path of countries with different socio-economic backgrounds and at different periods. This paper considers Indian districts as small open economies and compares their structural transformation path following the 1991 Indian trade reform and the spread of electricity use for irrigation in rural India. I use the interaction between the national ad-valorem tariff with the district level consumption share of each good that compose the consumption basket of rural households as a measure of the district level tariff. The identification strategy takes into account the potential endogeneity of the tariff measure and electricity for irrigation that are respectively instrumented by geographical accessibility variables of a given district and the district-level hydroelectric supply shift. Results suggest that a decline in tariffs' level, combined with the spread of electricity have a statistically significant effect on labor reallocation across sectors. The agricultural labor share decreases in districts that experienced a greater decrease in the tariff level and an increase of the labor share in other sectors of the economy. Consistent with the literature, the service sector (retail and education and public services) seems to absorb most of the labor from the agricultural sector.

4 INTRODUCTION

Developing countries are characterized by a substantial labor share in the agricultural sector. For most of them, agriculture can help reduce poverty in rural areas where most of poor individuals live. Nevertheless, value added per worker is four times higher in non-agricultural sector, while the labor share in the agricultural sector is much higher, attesting a misallocation of labor between sectors [Gollin et al., 2013]. We, therefore, assist to a labor reallocation during the development process of countries. In the specific case of India, the labor share in the agricultural sector was around 68.1% in the 1980s [Kochhar et al., 2006], and declined over the years to 48.9% in 2011-12 (International Labour Organization, 2016). [19] For this reason, it is important to pay attention to the structural transformation of developing countries and its determinants.

The literature theoretically established that the process of structural transformation of developing countries is highly influenced by the degree of openness of an economy. An increase in the agricultural productivity can lead to an industrialization or a specialization in the agricultural sector, depending on whether one is in a closed or open economy [Gollin et al., 2002, Matsuyama, 1992, Nurkse, 1953, Rosenstein-Rodan, 1943]. In a closed economy, the use of a new agricultural technology increases the output per worker and per unit of land. The surplus of workers in the agricultural sector is reallocated towards the manufacturing sector, leading to its expansion. In contrast, in the more realistic assumption of an open economy, we should observe a reallocation of the workforce towards the agricultural sector. The theory of comparative advantage tells us that an increase in agricultural productivity pushes the economy to specialize in the production of the agricultural good to the detriment of the industrial good ; leading to a progressive de-industrialization.

19. http://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-new$_d$elhi/documents/publication/wcms$_4$96510.pdf

However, there is little empirical evidence on how a change in trade policy will influence the relationship between agricultural productivity and structural transformation in developing countries. The non-existence of a completely open or closed economy makes this issue difficult to study and to empirically asses a situation of autarky or free trade. How important are the effects of trade policy on the relationship between agricultural productivity and structural transformation ? Does it favour an industrialisation or a de-industrialisation of developing countries ? In this paper, I exploit the timing of the Indian trade reform of 1991, with the spread of electricity adoption for irrigation, to assess the impact of trade liberalization on the relationship between agricultural productivity and labor reallocation among sectors in rural India. [20]

In 1991, the Indian government agreed to an adjustment program of the International Monetary Fund (IMF) in order to replace a protectionist economic regime with a competitive environment. A combination of factors pushed the Indian government to liberalize its trade policy : fiscal and balance of payment deficit ; the sudden increase of oil prices due to the Golf War in 1990 ; and political uncertainty due to the assassination of the Primer Minister Rajiv Gandhi [Edmonds et al., 2010a, Jain, 2016]. As in Edmonds et al. [2010a], I exploit differences in the national level of ad-valorem tariff on goods and services and the pre-reform heterogeneity in the consumption basket composition across Indian districts to construct a consumption-based tariff measure that reflects the impact of the level of tariffs applied on foreign goods on consumers' expenditures in a given district. Using this tariff measure I run a difference-in-difference approach comparing districts more and less exposed to liberalization. [21] Districts with a larger decreases in the consumer's perception of tariffs

20. Electricity for irrigation is considered as the agricultural technology. In India it is highly subsidized by the central and state governments for agricultural activities. According to Monari [2002], electricity subsidies for agriculture account for around 25% of India's fiscal deficit per year and is twice the annual public spending in health and rural development. Dubash [2007] argues that power consumption in agriculture was well above the one in the industrial sector by the mid-1980's

21. Districts are considered as small open economies.

and experiencing higher rate of electrification should exhibit faster labor reallocation towards the agricultural sector. This hypothesis aligns with Ricardo's theory of comparative advantage which states that an open economy specializes in the sector where it has a comparative advantage. The use of electricity in agriculture has the potential to increase agricultural productivity and to improve the comparative advantage of the agricultural sector.

The electrification rate of districts is not exogenous to economic activity [Allcott et al., 2016a, Gupta and Pelli, 2021, Rud, 2012]. The decision to develop electric infrastructure in a given area may depend on its economic dynamism and/or political climate. In the case of India, Modi [2005] argues that rural electrification is a campaign tool that politicians use to attract the vote of farmers. The level of electrification may, therefore, be correlated to unobserved factors that also determine the economic activity of districts. To tackle this endogeneity problem, I implement my identification strategy using an instrumental variable approach. I construct a Bartik instrument : the state level shifts in hydroelectric power supply weighted by the share of each district covered by surface water in 1970. [22] The hydroelectric supply shifts comes from Allcott et al. [2016a]. In their paper they use it to instrument electric shortages and show their effects on the performance of manufacturing plants in India. Given that hydroelectric power is the second source of energy in India behind coal, its supply shift may determine the quantity of electricity available in a given state. In the specific case of this paper, it may determine also the likelihood that farmers have access to electricity for irrigation. The share of a district covered by water in 1970 is used to transform the state level instrument to a district level one. In fact, water availability is crucial

22. Bartik [1991] was the first to develop such an instrumental variable by interacting the employment shares of a given industry located in a given area with the national employment growth of that industry. He used it to instrument the effect of employment on the real wage. Several other papers used the same strategy across many fields in economics : Blanchard et al. [1992], Bound and Holzer [2000], Autor and Duggan [2003], Chakravorty et al. [2014], Gupta and Pelli [2021]. See Goldsmith-Pinkham et al. [2020] for the latest development on the Bartik instrument.

for hydroelectricity production. It is therefore more plausible to install hydro-electric power plants in areas where water is available. And these areas are probably characterized by a higher electrification rate.

The second source of endogeneity is related to the tariff measure. In fact, the consumption based tariff measure may not reflect the actual impact of tariff on consumers' expenditures if for example a large share of households in a district consume locally produced goods and are not affected by the price of imports. In order to address this endogeneity problem, I regress the tariff measure on a series of geographical accessibility variables that determine to what extent a household has access to imported goods and services and I use the predicted values of the regression as instrument for the tariff measure (Frankel and Rose [2005] adopted a similar approach, with a gravity equation, to instrument for trade openness).

My results show opposing evidence to the theory of comparative advantage of Ricardo. I notice a labor reallocation at the individual level from the agricultural sector towards other sectors of the economy, following a combined effect of tariff reduction and an increase of electrification. In districts that experienced a greater decline of the tariff measure, the estimates imply a decrease of the agricultural labor share by 2.36 percentage point (p.p) and an increase in the manufacturing sector by 0.59 p.p. An increase of the labor share is also noticed in the retail (0.46 p.p), transport (0.15 p.p) and education & public services (1.33 p.p) sectors. The individual level analysis is confirmed by the household level with a decrease of the share of rural households which primary sector of activity is agriculture and an increase for other sectors of the economy. These results suggest that rural households in India might have taken advantage of the income effect of trade to invest in electricity connection for agriculture, allowing them to diversify their income sources. I also notice an increase of the number of days worked per week in the retail and education & public services sectors. Taken together, these findings show the benefits of trade openness for developing countries, given that the increase of their agricultural productivity will allow the development of their non-agricultural sectors.

I perform several tests to check the robustness of my results. First, I use an alternative measure of tariff by weighting the national ad-valorem tariff with the consumption share of goods produced in traded sectors only. [23] The reason is that the tariff of non-traded goods are put to zero in Edmonds et al. [2010a]'s database, making the tariff measure smaller than it should be in districts where individuals consume a large share of non-traded goods. Results that I obtain with the alternative measure of tariff are similar to my baseline results. Second, I test the hypothesis of exclusion restriction that suggests that the hydroelectric supply shifts influence labor reallocation between sectors only through a variation of the rural electrification rate. I run a reduced form regression of the outcome variables on the instrument for a sub-sample of districts that experienced a constant variation of their rural electrification rate between 1988 and 2000. I find no direct effect of the instrument on labor reallocation, suggesting that the exclusion restriction hypothesis is not violated. Finally, I perform a falsification test by randomizing the rural electrification rate and the tariff measure 1000 times and I run a regression each time. I, then, check the proportion of these regressions that display significant results. The objective of the test is to verify that my baseline results do not capture a simple correlation in the data.

In the literature, most studies develop a theoretical macroeconomic model and calibrate it to match the structural transformation of developing countries with an assumption of autarky or free trade. Stylised facts of structural transformation are also computed by researchers using long time series to show empirical evidence of the reallocation of econo-

23. The measure of tariff that I use in my baseline specification considers both traded and non-traded goods that compose the consumption basket of rural households. The alternative measure instead consider traded goods only.

mic activity across the broad sectors of agriculture, manufacturing and services [Echevarria, 1997, Gollin et al., 2007, Hayashi and Prescott, 2008, Teignier, 2018].[24] Conclusions of these models differ from each other depending on the assumptions that have been put forward. Their common point is mainly that the degree of openness is a determining factor in the process of structural transformation of an economy. For instance, Teignier [2018] in his two-sector growth model concludes that international trade plays a key role in the development process of low agricultural productivity countries by allowing them to import part of their food consumption.[25] This strategy (the imports of agricultural goods) allows them to focus on other sectors where they are more productive.

India offers an excellent setting for studying this issue because it allows to compare the structural transformation path of districts (considered as small open economies) which are subject to the same shock at the same time. This fills a gap left in the literature of structural transformation, as most papers compare the structural transformation path of countries with different socio-economic characteristics and often at different periods [Adamopoulos, 2011, Gollin et al., 2007, Teignier, 2018, Tombe, 2015]. In addition, developing countries are more and more oriented towards globalization and the formation of large regional economic partnership (African Continental Free Trade Area (AfCFTA), Trans-Pacific Partnership Agreement, EU-ACP Economic Partnership Agreements). The agricultural sector of these countries accounts for 40% to 70% of labour share (FAO, 2012). An improvement of the productivity in that sector can have major repercussions on the structural transformation of these economies.

24. See Herrendorf et al. [2014] and Gabardo et al. [2017] for a comprehensive review.

25. such as South Korea for the period 1963-2015 and Great Britain during the 19^{th} century.

The remainder of the paper is organized as follows. In section 5, I present the conceptual framework of my strategy. Section 6 presents an overview of electricity use for irrigation in India and why do I consider it as an agricultural technology, while section 7 presents the 1991 Indian trade reform. In section 8 and section 9 I present the data that I use for this study and my empirical strategy respectively. Section 10 shows my baseline results and some robustness checks. And Finally, section 11 presents the conclusion and a discussion on the implications of my results.

5 CONCEPTUAL FRAMEWORK

My empirical strategy is based on a model of structural transformation featuring trade barriers. This framework allows me to theoretically conceptualize the effect of tariff level applied to foreign products on the relationship between agricultural productivity and the structural transformation of developing countries. This effect operate through the income effect of trade following a variation of tariff.

Assume an economy which degree of openness depend on tariff level applied to foreign products. In this framework, one can suppose that agricultural productivity is endogenous and is determined by farmers' decision to adopt the agricultural technology. In fact, the investment required for the purchase of the technology may depend on the impact of tariffs on the consumption basket of households. A variation of tariff may generate an income effect. An increase of tariffs would increase the price of imported consumption and input goods that compose the consumption basket of households. In this case, one can expect households to devote a large share of their income to purchase consumption goods and decrease their savings that may allow them to invest in an agricultural technology. The majority of small-scale farmers, that are in the lowest part of the income distribution, would keep producing with a rudimentary technology, even if a better one is available. We would therefore expect no effect in terms of labor reallocation between sectors. In the case of tariff reduction, we may expect the opposite conclusion. Tariff reduction would decrease the price of consumption and input goods. Farmers may therefore cease the opportunity to increase their savings and invest in a new agricultural technology that enhances the productivity of factors of production (labor). In this case, a rational smallholder farmer employing family members for agricultural activities, may suggest the surplus of family labor to operate in other types of activity, such as the manufacturing or services sectors, in order to diversify households' income sources. An aggregation at the national level of these households' labor reallocation may generate the structural transformation of the economy.

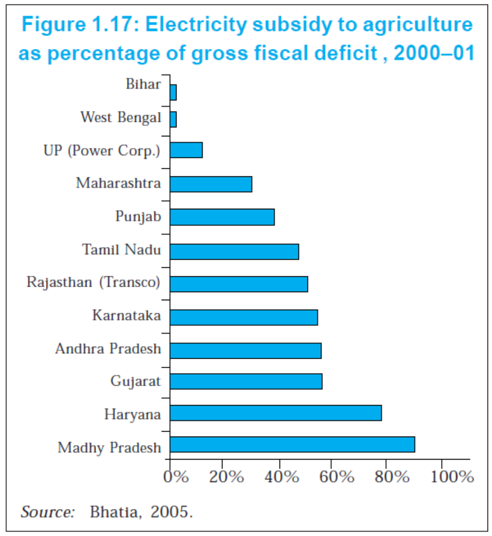
The effect of trade on the real income of households has been documented in the literature. According to Goldberg and Pavcnik [2007], there are three main channels through which trade liberalization may affect individuals : variation in labor income, variation in real income through changes in relative prices and variation in household production. Fajgelbaum and Khandelwal [2016] contribute to this literature by using data for a set of 40 countries and conclude that trade favor more poorer consumers. In fact, their spendings are mainly concentrated on traded sectors (food) that are more subject to international trade than non-traded sectors (services). The tariff level drop following trade openness allows, therefore, an increase of their real income. In the specific case of India, Topalova [2010] shows that the 1991 trade reform allowed a decrease of ”poverty incidence and depth” by about 15 percent suggesting that the trade reform might have had an income effect for Indian rural households. This income effect might have allowed them to invest in electricity adoption for agriculture and to diversify their economic activities.

6 OVERVIEW OF ELECTRICITY USE IN AGRICULTURE IN INDIA

In this section, I present the importance of electricity in the Indian agricultural sector and why I consider it as an agricultural technology. During the Bengal famine of 1943, around 2.7 million individuals died of starvation, due in large part to the poor agricultural performance [Sen et al., 1980]. Following this disaster, Indian government increased agricultural productivity in order to reach food self-sufficiency. A series of policies and programs have been adopted, including the Green revolution, launched in 1967/1968. This program included many components such as land reforms, the installation of irrigation infrastructures, the distribution of High Yielding Variety seeds (HYV) or electricity supply for irrigation [Foster and Rosenzweig, 2008].

6.1 Electricity for irrigation

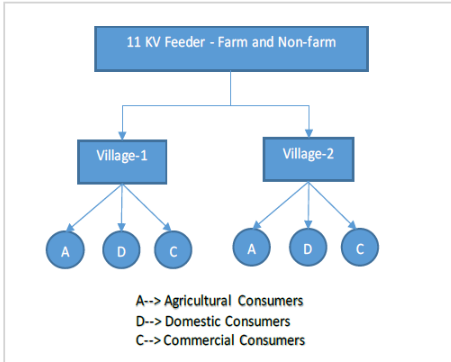
The Indian government aimed at providing electricity to farmers in rural areas for their irrigation needs and not to depend on rainfall for agricultural production. According to Ryan and Sudarshan [2020], India become the largest user of groundwater in the world, extracting more than the United States and China combined in a year. Briscoe and Malik [2006] argue that groundwater pumping uses around 40% of total electricity consumption. They also show that state governments make lot of financial efforts to provide electric power to farmers. In fact, electricity subsidies to agriculture constitute a large share of the gross fiscal deficit of Indian states. They are equal to about Rs 240 billion a year and is equivalent to 25 percent of India's fiscal deficit (Figure 3.1).



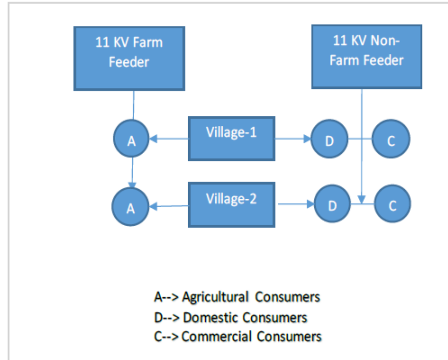
Note : Image from Briscoe and Malik [2006]

FIGURE 3.1 Electricity subsidy to agriculture as percentage of gross fiscal deficit , 2000 to 2001

In some states such as Gujarat, state governments go beyond providing electricity. They ensure that farmers have also an electricity of quality. For that, they put in place a feeder segregation program which objective is mainly to change the physical power supply infrastructure and to separate the non-agricultural and agricultural feeders [Chindarkara et al., 2017] (Figure 3.2). As a result, agriculture feeders provide 8 hours of high-quality electricity supply, allowing farmers to meet their irrigation needs. This evidence shows that electricity supply for irrigation is of central importance in India.



1a: Feeder system before JGY



1b: Feeder system after JGY

Source: Shah et al. (2008)

Note : Image from Chindarkara et al. [2017]

FIGURE 3.2 Gujarat's feeder segregation program : Jyotigram Yojana (JGY)

Even though electricity for irrigation is highly subsidized for farmers, it is not entirely free of cost. There exist mainly two ways for farmers to have access to groudwater for irrigation. First, at the beginning of the 1950s, the Indian government launched the construction of public tube wells that convey piped groundwater to fields with a flat tariff that is applied in order to facilitate the metering and fee collection and to recover part of the costs of electricity provision [Shah, Burke, Villholth, Angelica, Custodio, Daibes, Hoogesteger, Giordano, Girman, Gun, et al., 2007]. Second, rich farmers took advantage of the low installation costs to dig private tube wells in their fields [Chakravarti, 1973]. But the installation costs are high enough to prevent small-scale farmers to have their own tube wells and to depend on public and private well owners for groundwater irrigation. [26] The potential income effect of the 1991 trade reform might have been determinant in the adoption of electricity for irrigation by allowing small-scale farmers to be able to invest in this technology.

26. See Sidhu, Kandlikar, and Ramankutty [2020] for an overview of the evolution of the electricity demand for irrigation purposes in India.

6.2 Electricity as an agricultural technology

The consequences of electricity adoption in agriculture are numerous. They fit into the literature on agricultural technology adoption and labor reallocation in developing countries [Bustos et al., 2016, Emerick, 2018, Foster and Rosenzweig, 2004, 2007, Hornbeck and Keskin, 2015]. As suggested by Banerji, Meenakshi, and Khanna [2012], adequate electricity supply to farmers has the potential to increase agricultural yields with less resources. Suppose a smallholder farmer employing family members for farming activities. Without electricity, she relies on family members for watering plants. With electricity, less labor may be needed because the heavy duty is done by the electric pump. In this case, the surplus of family labor may be reallocated towards other sectors of the economy to diversify household income sources. In addition to allowing an increase in production with less labor, the use of electricity also allows farmers to produce water intensive crops such as sugar, rice or sorghum. Badiani and Jessoe [2013] argue that electricity subsidies for groundwater extraction in India allow the increase of the output of mainly water intensive crops.

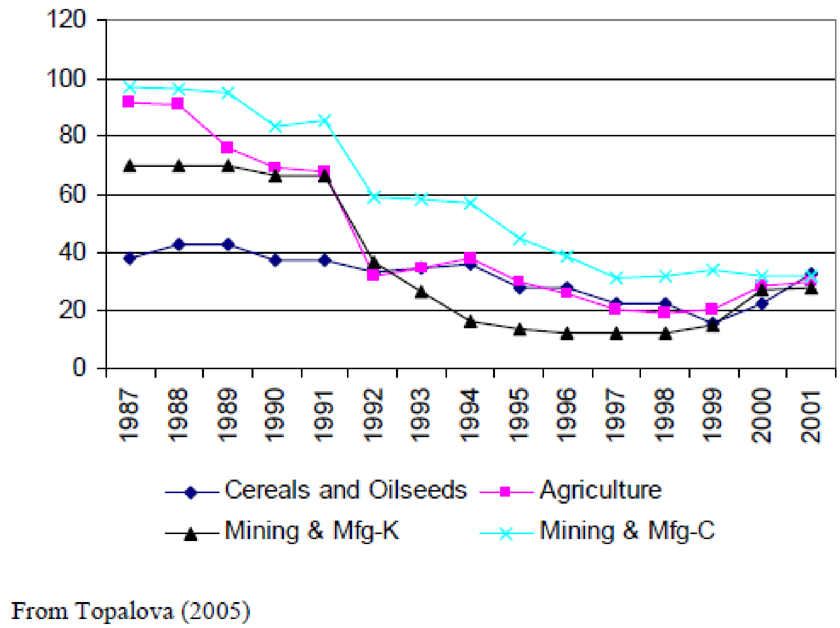
7 THE INDIAN TRADE REFORM

Studying the effect of trade openness on the relationship between agricultural productivity and structural transformation is not easy, mainly because of the impossibility to compare a situation of autarky with a situation of free trade. India offers a good setting for this kind of study with its 1991 trade reform. Before 1991, the country was characterized by a restrictive trade regime with import protection, government ownership of heavy industries or complex industrial licensing requirement [Cerra and Saxena, 2002]. "India's trade restrictions were among the most severe in the world, and utilized a variety of tools : high tariff and nontariff barriers, a complex import licensing system, an actual user requirement that prohibited imports by intermediaries, restriction of certain imports to the public sector ("canalization"), and government purchase preferences for domestic producers, among others." (Topalova, 2010, p. 4).

A combination of factors pushed the government to adopt the 1991 trade reform imposed by the International Monetary Fund (IMF). In the 1980's, the fiscal and balance of payment deficit put the country in an economic crises. The situation was exacerbated by the sudden increase in oil prices due to the Gulf War in 1990, political uncertainty with the assassination of Rajiv Gandhi, the leader of the Congress Party and undermined investor confidence [Topalova, 2010]. In August 1991, the IMF accepted to help the country resolve its macroeconomic problems with the condition to operate structural reforms and macroeconomic stabilization of the Indian economy. The reform has affected many levers of the Indian economy, including the liberalization of the economy with the reduction of taxes on goods and services. [27]

All sectors of the economy have been affected by the tariff reform. Figure 3.3 shows the evolution of tariff in manufacturing, cereals and oilseeds, agriculture (other than cereals and oilseeds) and mining. Figure 3.3 also shows an increase of tariff in some sectors after 1997. According to Edmonds et al. [2010a], this increase may reflect various political economy factors, causing the non-exogeneity of tariff reductions for the period after 1997. They, therefore, restrict tariff level for the period 1987-1997.

27. For more details about the Indian trade reform see Edmonds et al. [2010a], Topalova [2010] and Varshney [1998].

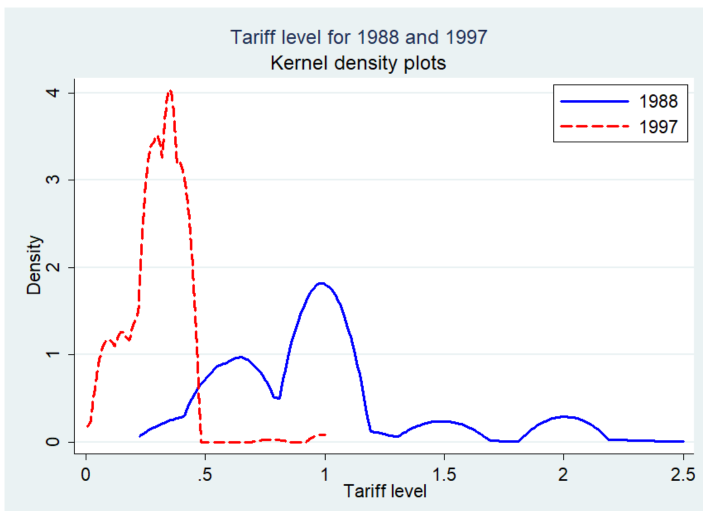


Note : Graph from Edmonds et al. [2010a].

FIGURE 3.3 Tariffs by Industry

Topalova [2010] presents several pieces of evidence proving the exogeneity of the tariff reductions between 1987 and 1997. First, she argues that the reform was unanticipated and did not influence individuals' decisions and the industrial composition of districts prior to 1991. Goyal [1996] describes the reform as a "shock therapy" in the sense that it was a condition of the IMF following India's currency crisis and was quickly adopted in order to avoid political resistance. Second, Topalova [2010] investigates whether there has been a political will of the Indian authorities to favor certain industries to face international competition by changing their trade protection. She finds that there was a uniform movements in tariffs for almost 5000 items for the period 1992-1997. But for the period 1997-2001, she argues that "tariff movements were not as uniform" and suspects political or economic intentions in their evolution. Third, Topalova and Khandelwal [2011] test whether Indian policy-makers adjusted tariff according to the productivity of vulnerable industries in order to protect them from foreign competition. They find no correlation between tariff level and productivity for the period 1989-1996, but observe a significantly negative correlation for the period 1997-2001. That is the reason why, Topalova [2010], Edmonds et al. [2010a] and Topalova and Khandelwal [2011] focus only on the 1987-1997 period for their studies.

As shown by Figure 3.4, the nominal ad-valorem tariff level on traded goods was high prior to 1991, with an average of 97%, and declined to an average of 30% in 1997. [28] This shows how India opened up to international trade following the trade reform. I will use this variation to study how labor reallocates between sectors following an increase in agricultural productivity.



Note : Author's graphic from Topalova [2010]'s database. The blue and red line represent the distribution of the nominal ad-valorem tariff for traded goods in 1988 and 1997 respectively. The tariff of non-traded goods such as service, trade, transportation and construction and also cereals and oilseeds are put to zero for the period 1987-1997 in Topalova [2010]'s database. For this reason, these goods are not considered in this graph.

FIGURE 3.4 Indian nominal tariff level comparison between 1988 and 1997

28. The tariff of non-traded goods such as service, trade, transportation and construction and also cereals and oilseeds are put to zero for the period 1987-1997 in Topalova [2010]'s database.

8 DATA

I combine data on the employment and unemployment of rural individuals, their consumption and expenditure and district level data to have a repeated-cross section for years 1988 and 2000. Data are from several sources : the rural sample of the National Sample Survey (NSS) of India, Edmonds et al. [2010a], Allcott et al. [2016a] and the PRIO-GRID. [29]

8.1 Individual and household data

The National Sample Survey (NSS) is a repeated cross section database collected periodically at the individual and household level by the Ministry of Statistics and Program Implementation of India and is representative at the district level. I use rounds 43 (1987-1988) and 55 (1999-2000) of the "employment and unemployment survey". [30] The data contain information on the principal and subsidiary activity status of individuals and on the number of days worked in each activity. NSS also includes detailed informations on the industry in which individuals work at the time of the survey, which is essential in order to study structural transformation. The database provides a five digit code indicating the industry in which individuals operate. The first digit indicates broad sectors : agriculture, manufacturing, construction, retail, transportation and education & public services. [31] The first part of Table 3.1 shows summary statistics on the employment of rural individuals in the NSS. Focusing first on the employment share in each sector, I notice a fall of the workforce in the agricultural sector from 74% to 72% between 1988 and 2000. The difference is significant at 1% level. The manufacturing, retail and transport sectors seem to be the sectors that may have absorbed much of the workforce that have left the agricultural sector. The labor share

29. `https://www.prio.org/Data/PRIO-GRID/`

30. NSS 50 (1994) would be closer to the trade reform of 1991, yet it does not allow to identify districts of residence. For this reason, I use NSS 55 (2000) as the post-reform period.

31. I regroup the education and public services sectors because education is primarily provided by public schools in India.

in 1988 in these sectors was 6%, 4% and 1% respectively. And it increased to 7%, 5% and 2% in 2000. For all these sectors, the difference of labor share for the two years is significant at 1% level. Table 3.1 also provides information on the evolution of the number of days worked per week by individuals in each sector. I mainly notice a decrease of the variable for all sectors except the transport sector.

A different survey of the NSS provides detailed information on the consumption and expenditure of households that are collected the same years as the employment and unemployment survey : 1988 and 2000. I use the information on the main source of lighting for rural households, to determine which one is connected to the electric grid. This allows me to estimate the rural electrification rate which I use as an indicator of farmers' accessibility to electricity for their irrigation needs. Table 3.1 shows that the rural electrification rate was about 30% in 1988 and the number rose to 55% by 2000.

Table 3.1: Descriptive Statistics

	1988		2000		t-test		Observations
	Mean (1)	SD (2)	Mean (3)	SD (4)	Diff (5)	t-Stat (6)	(1988 + 2000) (7)
Panel A :Employment Share							
Agricultural sector*	0.74	0.44	0.72	0.45	0.02	6.72	445423
Numb. of days worked in agriculture	2.19	2.9	1.66	2.66	0.54	58.11	445423
Manufacturing sector*	0.06	0.24	0.07	0.25	-0.005	6.10	445423
Numb. of days worked in manufacturing	0.18	1.04	0.15	0.95	0.03	8.85	445423
Construction sector*	0.03	0.18	0.03	0.18	-0.002	0.34	445423
Numb. of days worked in construction	0.09	0.69	0.07	0.62	0.02	9.03	445423
Retail sector*	0.04	0.2	0.05	0.22	-0.01	-14.76	445423
Numb. of days worked in retail	0.13	0.9	0.12	0.87	0.01	3.42	445423
Transport sector*	0.01	0.11	0.02	0.14	-0.008	-19.33	445423
Numb. of days worked in transport	0.04	0.47	0.05	0.54	-0.01	-6.13	445423
Education & Public sector*	0.06	0.24	0.03	0.18	0.03	42.43	445423
Numb. of days worked in educ. & Public sector	0.18	0.2	0.08	0.7	0.1	32.59	445423
Panel B : Electrification rate, Hydroelectricity & tariff measure							
Rural electrification rate	0.30	0.23	0.55	0.3	-0.25	-301.47	652
Hydroelectricity supply	0.34	0.34	0.26	0.27	0.08	77.38	48
Hydroelectricity * $Water_{d,s,1970}$	0.61	2.39	0.46	1.78	0.15	20.9	652
National level of ad-valorem Tariff (traded goods)	0.97	0.41	0.3	0.13	.67	27.63	
District level tariff ($tariff_{d,t}$)	0.57	0.08	0.18	0.04	0.39	1746.98	652
Panel C : Climate Data							
Rainfall (m^3)	3.98	2.53	3.16	1.97	0.82	105.8	652
Temperature (°C)	25.72	3.11	25.41	3.56	0.3	28.24	652
Panel D : Socio-Demographic Variables							
Banks per capita	0.06	0.02	0.08	0.03	-0.01	-168.9	652
Poverty Gap	0.08	0.06	0.05	0.04	0.04	211.49	652

Notes : Variables with a (*) indicates dummy variables. For instance, *Agricultural sector* is equal to one for individuals that have their primary activity in the agricultural sector. The electrification rate, the tariff measure and socio-demographic variables are computed at the district level. The hydroelectricity supply shift is at the state level. In total, I have 326 districts and 24 states for each year (1988 and 2000). Variables related to the employment share are computed at the individual level. Column (5) and (6) show if the mean difference between 1988 and 2000 is significant. $Water_{d,s,1970}$ is the percentage area of the district covered by water area in 1970.

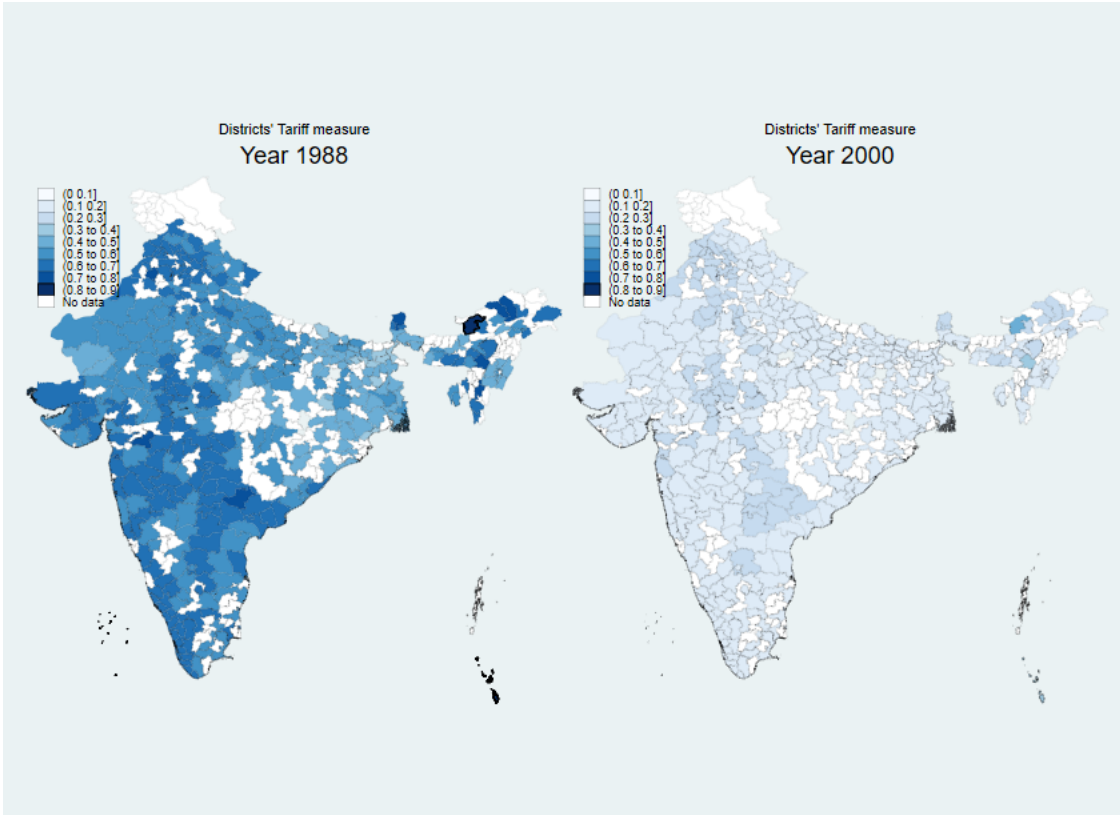
8.2 Tariff measure : Consumer's perception of tariff

Following Edmonds et al. [2010a], I measure tariff protection at the district level by interacting the share of total expenditure on good p in district d, with the tariff applied to good p. The district tariff measure at time t is computed as follows :

$$tariff_{d,t} = \sum_{p} consshare_{p,d,1987} * tariff_{p,t} \tag{3.1}$$

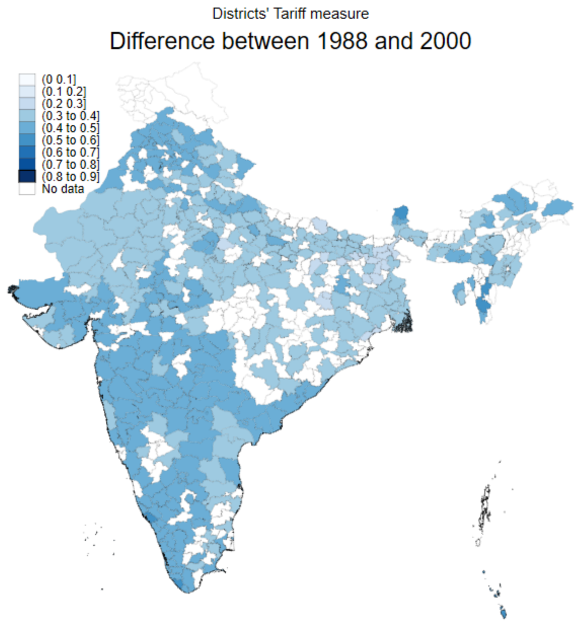
and is "a measure of the consumer's perception of tariffs in a given district". $tariff_{p,t}$ is the tariff on good p at time t, constant across district and $consshare_{p,d,1987}$ is the share of total expenditures in district d in 1987 spent on good p. This measure of tariff is computed using district-specific consumption weights based on the consumption basket of Indian households prior to the trade reform of 1991. [32] Thus, changes in the consumption basket over time due to the tariff reform do not influence the measure of tariff [Edmonds et al., 2010a]. Figure 3.5 shows the heterogeneity of districts' exposure to tariffs for 1988 and 2000, while Figure 3.6 shows the difference between the two years.

32. Edmonds et al. [2010a] say that they use a large set of consumption goods purchased by Indian households, with information on home production and an array of food and non-food goods, to construct district specific consumption weights for goods.



Note : Author's graph from Edmonds and al. (2010) data.

FIGURE 3.5 Heterogeneity in districts' tariff for years 1988 and 2000



Note : Author's graph from Edmonds and al. (2010) data. The graph displays for a specific district, the difference of the Tariff measure between 1988 and 2000.

FIGURE 3.6 Difference in districts' tariff between 1988 and 2000

As explained in Edmonds et al. [2010a], the trade literature generally use an employment-based tariff measure to study the impact of trade reforms on socio-economic outcomes. My main outcome variables are related to the employment share of the different sectors of the Indian economy. For this reason, I consider an alternative channel, with a consumption-based tariff measure. Table 3.1 shows that $tariff_{d,t}$ decrease from 57% in average in 1988 to 18% in 2000, indicating that the tariff measure declined in average by around 39 percentage point. This reflects the fact that the trade reform may have had an income effect for Indian households with the decrease of consumption and intermediate input prices [Edmonds et al., 2010a]. This income effect is a possible channel through which trade reform can lead to the adoption of new agricultural technologies by rural households and trigger labor reallocation between sectors. In fact, the income effect generated by the reform may allow agricultural households to increase their savings and invest in a new agricultural technology (like electricity) that enhances their production.

8.3 Hydroelectricity supply shifts

I use the instrumental variable developed by Allcott et al. [2016a] : the predicted hydroelectricity generation as a share of the predicted electricity demand at the state level. It is constructed for the period 1992-2010. [33] Panel B of Table 3.1 shows that hydroelectricity supply shift represented 34% of the electricity demand in 1992 and declined to 26% in 2000. Hydro power generation represents the second source of energy in India behind coal.

8.4 The PRIO-GRID

I finally use the PRIO-GRID database that is a georeferenced dataset that comes with cell-specific information on a large selection of political, economic, demographic, environmental and conflict variables covering the period 1946 to 2013. Each cell has a resolution of 0.5 x 0.5 decimal degrees, that corresponds to a cell of roughly 55 x 55 kilometers at the Equator (3025 square kilometers area). Data that I need from this dataset are climate data such as the temperature (1948-2014) and rainfall (1979-2014) and also data related to the geographical accessibility variables of districts.

9 EMPIRICAL STRATEGY

My empirical specification seeks to estimate how trade policy influences the relationship between agricultural productivity and the sectoral reallocation of labor. Denoting $Y_{i,d,t}$ as one of the outcome variables described above at the household/individual level (i), in district d and year t (1988 and 2000), the equation of interest is as follows :

33. I will give more details about the construction of this variable in section 6.

$$Y_{i,d,t} = \beta_0 E_{d,t} + \beta_1 tariff_{d,t} + \beta_2 E_{d,t} * tariff_{d,t} + \beta_3 X_{d,t} + \alpha_d + \gamma_t + \epsilon_{i,d,t} \tag{3.2}$$

where $E_{d,t}$ is the rural electrification rate. It represents the proxy of electrification in a district. The interaction between $E_{d,t}$ and $tariff_{d,t}$ will capture the combined effect of agricultural productivity and trade openness on labor reallocation between sectors in rural India. $X_{d,t}$ is a matrix of covariates at the district level : the average temperature and rainfall, the number of banks per 1000 people and the poverty gap. α_d and γ_t are respectively district and time fixed-effects. $\epsilon_{i,d,t}$ is the error term. Standard errors are clustered at the village level. Rural electrification programs in India have the objective to provide electricity to villages. [34] Thus, individuals living in the same village have access to similar quantity and quality of electricity. Clustering at the village level allows to take into account these facts.

An OLS estimation of equation (3.2) is with no doubt inconsistent because of the endogeneity of the rural electrification of districts and the tariff measure. I present below the source of endogeneity of both variable and how I address the problem.

9.1 Endogeneity of the electrification level of districts

The problem when estimating equation (3.2) lies in the fact that it is difficult to isolate the impact of investment in electrification on economic outcomes. The main reason is that electrification may depend on unobservables that may confound its effective effect [See for instance Allcott et al., 2016a, Gupta and Pelli, 2021, Rud, 2012]. For instance, the economic dynamism and/or the political climate of an area may determine its level of

34. `http://www.indiaenergy.gov.in/doc/Expert-view/Rural%20Electrification%20India.pdf`

electrification compared to others. For example, the population density, the proximity to a major city or the level of development of villages may influence their probability to be connected to the electric grid.

To address this endogeneity problem, I consider the instrumental variable developed by Allcott et al. [2016a], that is the predicted hydroelectricity generation that vary from year to year due to the availability of water in the reservoir of hydroelectric power plants. The Indian government argues that hydro power is more reliable and affordable than fossil fuels, because it is a renewable power source.[35] Chan et al. [2014] show that Indian coal power plants were offline 28% of the time between the period 1994-2009. In this case, the advantage using the hydroelectric supply shifts as an instrument is that it captures both the quantity and quality of electricity supply compared to thermal power. In addition, water availability and the topology of the area are crucial for hydroelectricity generation. Thus, the installation of hydroelectric power plants in a given area may depend mainly on these two determinants and not on economic activities. Applied to the context of this paper, I can say that a farmer will use electricity for irrigation purposes if and only if it is available in his area. The amount of electricity available, as well as the probability for a household to have access to electricity, may be related to the supply shift from hydroelectricity generation in his home state.

35. https://www.energy.gov/eere/water/benefits-hydropower

Allcott et al. [2016a]'s instrument ($Hydro_{s,t}$) is computed at the state level and represents the share of predicted electricity demand covered by state s predicted Hydroelectricity generation. [36] In their paper, they use it to instrument electricity shortages. [37]

A threat to the exclusion restriction hypothesis is that there is the possibility for a hydroelectric central to store water in reservoir for future years if the demand of electricity of agricultural or non-agricultural sectors is low in the current year. Actual hydroelectricity generation would, in this case, be related to the economic activity. For this reason, authors use predicted hydroelectricity supply and predicted electricity demand to compute the share of state s electricity demand satisfied by hydroelectricity supply and use it as instrumental variable for electricity shortages in India.

Since I intend to do a district-level analysis, I use a Bartik instrument by interacting the hydroelectricity supply shifts by the share of a district covered by water in 1970 ($Water_{d,s,1970}$). This strategy allows to transform the state level instrument to a district level one.

$$Hydro_{d,s,t} = Hydro_{s,t} * Water_{d,s,1970} \tag{3.5}$$

36.

$$Hydro_{s,t} = \frac{H_{s,t}}{\tilde{Q}_{s,t}} \tag{3.3}$$

Where $H_{s,t}$ is the predicted state-level hydroelectricity generation computed by using reservoir inflows and generation from "run-of-river" hydro plants that have no reservoirs to store water. And $\tilde{Q}_{s,t}$ represents the predicted electricity demand computed as follows :

$$\tilde{Q}_{s,t} = \sum_{r \neq s} Q_{r,t}. \sum_{y=1992}^{2010} \frac{Q_{s,y}}{\sum_{r \neq s} Q_{r,y}} \tag{3.4}$$

By dividing predicted hydroelectricity generation by the predicted electricity demand we obtain the relative share of hydro generation across states.

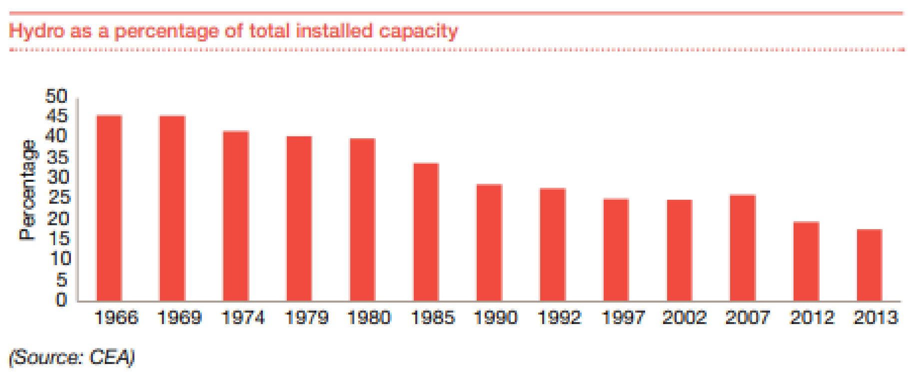
37. Allcott et al. [2016a] are interested in estimating how variation in electricity shortages affects the performances of manufacturing plants in India. The possibility that economic growth may increase electricity demand and therefore electricity shortage, leads to an endogeneity problem coming from reverse causality. To address the problem, the authors use the predicted supply shift of hydroelectricity generation as an instrument for shortages.

The idea is that the quantity and quality of electricity from hydroelectric plants available in a district is related to its "initial" level of water available. In fact, India started hydroelectric power development at the beginning of the 20^{th} century. After the independence in 1947, the country paid considerable attention to the generation of power and massive investments have been made for hydroelectricity generation as it is the most economic and preferred source of energy [Saxena and Kumar, 2010]. The decision to install hydro plants in a given area may have been influenced by the availability of water, given that the main input for hydro generation is water and also the topology of the district. [38] It is therefore more plausible to install hydro-electric power plants in areas where water is available. Thus, interacting the hydro supply shift by the share of a district covered by water in 1970 is equivalent to weighting it by the the likelihood to invest in hydroelectric facilities in a given area during the post-independence period. In addition to that, Rud [2012] explains that the uneven availability of groundwater across states - as measured by the thickness of the watertable - was an important determinant of electrification, as the High Yield Variety (HYV) seeds introduced during the Green Revolution depended, among other factors, on the provision of timely irrigation. A district characterized by a high share of water area in 1970 may have a greater probability to have a high electrification rate and to use more effectively a positive hydro supply shifts. [39]

38. Allcott et al. [2016a] explain that the "annual output of hydro plants depends primarily on a water availability constraint determined by rainfall at higher elevations (or snowfall in the Himalayan states)".

39. The necessary condition for doing this transformation is to make sure that the new instrumental variable at the district level ($Hydro_{d,t}$) is still valid ; that is to say $E(Hydro_{d,t} * \epsilon_{d,t}) = E(Hydro_{s,t} * Water_{d,s,1970} * \epsilon_{d,t}) = 0$. Taking the limit of this expression, we can have : $\lim\limits_{S,D \to \infty} \frac{1}{D*S} \sum_D \sum_S (Hydro_{s,t} * Water_{d,s,1970} * \epsilon_{d,t}) = 0$. $\lim\limits_{S,D \to \infty} \frac{1}{D*S} \sum_S Hydro_{s,t} \sum_D (Water_{d,s,1970} * \epsilon_{d,t}) = 0$. And this condition is verified if and only if $\lim\limits_{D \to \infty} \frac{1}{D} \sum_D (Water_{d,s,1970} * \epsilon_{d,t}) = 0$ or $E(Water_{d,s,1970} * \epsilon_{d,t}) = 0$. This condition holds if and only if shocks related to the different sectors of the Indian economy in 1988 and 2000 are not correlated to the share of a district covered by water in 1970.

A problem with this instrument, that is worth noting, is that it is computed for the period 1992-2010. This study uses socio-economic data of 1988 and 2000. I, therefore, assign the instrumental variable data of 1992 to the socio-economic data of 1988. Figure 3.7 shows that the share of hydroelectricity with respect to the total installed capacity is between 30% and 35% during the period 1985-1992. And in Table 3.1, I have a share of the hydro supply of around 34% in 1988. Given this small variation of the share of hydro supply between 1985 and 1992, I assume that its level in 1988 is similar to the one in 1992.



Note : Image from PricewaterhouseCoopers (2014)

FIGURE 3.7 Hydro as percentage of installed capacity in India

9.2 Endogeneity of districts' exposure to tariffs

A series of factors may determine the degree of openness of an economy. For instance, the influence of interest groups on politicians may affect their political and economic decisions regarding the trade policy adopted [Grossman and Helpman, 2002]. For this reason, trade openness is considered as endogenous in the literature. As said in section 3.6, Topalova [2010] presented several features of the 1991 Indian trade reform that suggest that it was exogenous to economic and political activities during the period 1988-1997. The endogeneity problems that I address here are related to the construction of the tariff measure at the district level.

Recall that districts' exposure to tariff is the interaction between the share of total expenditures in district d in 1987 spent on good p and the nominal ad-valorem tariff on good p. In Topalova [2010]'s database, tariff of non-traded sectors, such as services, transportation, trade and construction and also the cultivation of cereals and oilseeds, are put to zero. [40] The tariff measure ($tariff_{d,t}$) is therefore sensitive to the share of non-traded goods in a households' consumption basket. Given that the majority of individuals in poorest districts consume a large share of cereals, the tariff measure is, therefore, highly influenced by the level of poverty of districts. Putting the tariff of cereals to zero may lower the tariff measure in poorest districts. [41] This underestimation may cause a potential endogeneity in the tariff measure. Controlling for districts poverty level may alleviate this concern. In particular, I use the poverty gap, that is the proportion of the population below the poverty line, as a covariate in my specifications.

Second, districts' exposure to tariff is a consumption based measure that depends on the structure of the consumption basket of Indian households. For example, the level of tariff may have no effect on households/individuals that mainly consume local goods produced in India. In this case, the tariff measure might not reflect the actual consumers' perception of tariff if a large share of households in the district mainly consume local goods. This raises a potential problem of measurement error that causes the endogeneity of the tariff measure. To address this problem, I use the geographical accessibility of certain areas as a determinant of the likelihood for residents to have access to imported goods. In fact, areas difficult to access are more likely to have a lack of access on goods and services. I, therefore,

40. Attributing zero or infinite tariff to non-traded sectors does not matter since they do not change over time [Topalova, 2010]. Moreover, only government agencies were allowed to import cereals and oilseeds and no change in their tariff rates was observed. That is why their tariffs are also put to zero.

41. According to the Reserve Bank of India (2013), poorest states are Chhattisgarh, Jharkhand, Manipur, Bihar, Odisha and Madhya Pradesh and are mainly located at the north-east of the country.

construct an instrumental variable that determine whether or not imported goods are easily accessible. I regress the tariff measure on a series of geographical accessibility variables and use the predicted values of the regression as an instrument for the tariff measure. [42]

$$tariff_{d,t} = \sum_i \beta_i GA_{d,t} + \mu_{d,t} \tag{3.6}$$

Where $GA_{d,t}$ is a vector of geographical accessibility variables related to the travel time to the nearest major city, the proportion of mountainous terrain within the district, the distance in kilometer to the border of the nearest land-contiguous neighboring country, the distance to the border of the nearest neighboring country, regardless of whether the nearest country is located across international waters, the shortest straight-line distance to international waters, the distance to the national capital, the annual average precipitation and temperature. [43] The predicted tariff measure is a valid instrument if it determines the share of local and imported goods that compose the consumption basket of Indian households and affects the outcomes of the different sectors of the Indian economy only through the tariff measure.

First stage

The first stage specifications take the following form :

$$W_{d,t} = \gamma_1 Hydro_{d,t} + \gamma_2 Hydro_{s,t} + \gamma_3 PredictedTariff_{d,t}$$

42. A similar strategy is adopted in the trade literature by Frankel and Rose [2005] who use exogenous determinants of trade - geographical variables from the gravity model - as instruments to isolate the effect of openness. They regress trade openness on geographic variables such as physical distance, population size, land area as well as dummy variables indicating common borders, linguistic links, and landlocked status and use the prediction of the regression as instrument for openness.

43. Equation 3.6 is estimated by OLS. All geographical accessibility variables and their definition come from the PRIO-GRID.

$$+ \gamma_4 Hydro_{d,t} * PredictedTariff_{d,t} + \gamma_5 X_{d,t} + \alpha_d + \alpha_t + \lambda_{d,t} \quad (3.7)$$

where $W_{d,t}$ represents one of the endogenous variables in equation (3.2) : the rural electrification rate ($E_{d,t}$), the tariff measure ($tariff_{d,t}$) or their interaction ($E_{d,t} * tariff_{d,t}$). $Hydro_{d,t}$ is the district level instrument obtained by interacting Allcott et al. [2016a]'s instrument ($Hydro_{s,t}$) and the share of a district covered by water in 1970 ($Water_{d,s,1970}$). $PredictedTariff_{d,t}$ is the predicted values obtained from the regression of equation (3.6). $X_{d,t}$ is the vector of district level controls used in equation (3.2). Finally, α_d and α_t are district and year fixed-effects respectively and $\lambda_{d,t}$ is the error term, clustered at the village level.

Table 3.2 shows the results of the first stage specifications. Column (1) considers the rural electrification rate as the dependant variable. Column (2) focuses on the interaction between the rural electrification rate and the tariff measure. Finally, column (3) is interested on the tariff measure. Column (1) shows that the district level instrument has a positive and statistically significant effect at 1% level on the rural electrification rate of districts. The size of the estimate suggests that a one unit shift in Hydro supply, combined with a higher share of the district area covered by water in 1970, increases rural electrification by 0.151 percentage point. The sign of the coefficient suggests that a positive variation of the hydro supply has a positive effect on rural electrification. The state level instrument ($Hydro_{s,t}$) is negative but has to be interpreted as a marginal effect with coefficients on the interaction terms. A district with an average share of water area and predicted tariff, experience an increase of the rural electrification rate by 0.052 percentage point, following a one unit increase of the hydro supply shift. [44]

44. The share of a district area covered by water in 1970 range from 0 to 39.8%, with an average of 1.73%. The predicted tariff range from 0.375 to 0.677, with an average of 0.452. Districts that have the maximum share of water area and predicted tariff have an increase in rural electrification rate by 5.74 percentage point. In contrast, district that are at the minimum, experience a decrease of the rural electrification rate by 0.08 percentage point.

Column (2) presents the first stage with the interaction between electrification and tariff as the dependant variable. There is no economic importance in interpreting the results of this column. Statistically, I notice a significant effect of the of the interaction of the two instruments on the interaction of the two endogenous variables.

Column (3) shows that there is a strong positive correlation between the measure of districts' exposure to tariff and the predicted tariff and is equal to 0.668. [45]

Given that the error term in each specification is clustered at the village level, I present the Kleibergen-Paap rk Wald F statistic. For each specification, they suggest that the instrumental variables are relevant.

45. Frankel and Rose [2005] have a correlation between trade openness and their predicted instrument of 0.72 for a set of country. They use the ratio of trade to output as a measure of openness.

TABLE 3.2:
First stage

	Rural elec. rate	Rural elec. rate * Tariff	Tariff
	(1)	(2)	(3)
$Hydro_{d,t}$ ♠	0.151***	-0.014*	-0.052***
	(0.019)	(0.007)	(0.006)
$Hydro_{s,t}$ ♣	-0.08***	0.009**	0.004
	(0.011)	(0.004)	(0.003)
$Hydro_{d,t} * PredictedTariff$	-0.284***	0.036**	0.104***
	(0.038)	(0.015)	(0.013)
Predicted tariff	1.382***	0.911***	0.668***
	(0.227)	(0.133)	(0.067)
Banks per 1000 people	1.339***	0.091	0.433***
	(0.123)	(0.072)	(0.046)
Temperature	-0.014***	-0.002**	-0.0008
	(0.002)	(0.0009)	(0.0005)
Rainfall	0.003***	0.001**	-0.004***
	(0.001)	(0.0007)	(0.0004)
District poverty gap	-0.851***	-0.512***	-0.269***
	(0.043)	(0.023)	(0.015)
District and year F-E	Yes	Yes	Yes
Observations	445421	445421	445421
F-Stat	62.90	132.36	80.56
Kleibergen-Paap rk Wald F statistic	25.08		

Notes : Standard errors in parentheses are clustered at the first-stage unit (fsu) level that corresponds to villages in rural areas. [a] Predicted tariff represents the predict of the regression of the tariff measure on geographical accessibility variables and is the instrument of the tariff measure. ♠ is the district level instrument for the Rural elec. rate : $Hydro_{s,t} * Water_{d,s,1970}$. ♣ is Allcott et al. [2016a] instrument at the state level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

a. In the National Sample Survey there is three stage in the sampling design : the first-stage unit (fsu) that is villages and blocks for rural and urban areas, an intermediate stage of sampling that is the selection of two hamlet-groups (hgs)/ sub-blocks (sbs) from each rural/ urban FSU and the ultimate stage units (USU) that are households in both the sectors. I am clustering my standard errors at the FSU level.

10 RESULTS

First, I present the main results on how the combined effect of a tariff decline and electricity adoption affects labor reallocation across sectors. This analysis is done at the individual and household level. I also show evidence on the number of days worked in each sector for individuals. Second, I present a robustness check using an alternative tariff measure.

I, also, test the hypothesis of exclusion restriction of the hydroelectricity supply shifts as an instrumental variable. Finally, I present a falsification test that consist in randomizing the rural electrification rate and the tariff measure over the sample.

10.1 Effect on labor reallocation

Table 3.3 reports results on labor reallocation in rural areas. Table 3.4 focuses on results on the primary activity of rural households. And table 3.5 presents results on the number of days worked in each sector. In all specifications, standard errors are clustered at the village level. I include also district and year fixed-effects, as well as controls at the district level.

Individuals primary sector

Column (1) of Table 3.3 shows a negative direct effect of electricity on the labor share of the agricultural sector. A one percentage point increase in the rural electrification rate is associated with a decrease of the labor share in the agricultural sector by 1.33 percentage point. Because of the interaction with the tariff measure, the effect of electricity has to be interpreted as a marginal effect and becomes $-1.33 + 2.02 * \Delta Tariff$.[46] Districts in which tariff decrease more experience a larger decrease of the labor share in the agricultural sector, following an increase of electrification between 1988 and 2000. The marginal effect suggests a decrease of the agricultural labor share by 2.36 p.p, 2.12 p.p and 1.87 p.p, for districts experiencing the largest, average and lowest tariff variation respectively.[47] Figure 3.8 shows that the average rural electrification rate rose by around 25 p.p between 1988 and 2000. This implies a decrease of the agricultural labor share by 53% for districts at the average decline of tariff; equivalent to a decrease from 74% to 48% over the 12-year period.

46. Given that I run a fixed effect regression with two periods, the specification is similar to a first difference regression.

47. The tariff measure decline range from 0.515 to 0.272 and its average is 0.392.

I observe the opposite effect in the other sectors of the economy, with an increasing of their labor share in districts that experience a larger decline of tariff perception. For the manufacturing sector, in column (2), the direct effect of electricity use for irrigation implies that a 1 percentage point increase of the rural electrification rate can be associated to an increase of the employment share by 0.24%. By taking into account the interaction with the tariff measure, the effect becomes $0.24 - 0.69 * \Delta Tariff$. The marginal effect implies an increase of the labor share by 0.43 p.p, 0.51 p.p and 0.59 p.p for districts at the lowest, average and highest decline of tariff respectively. Results are qualitatively similar for the labor share in the retail, transport and the education & public services sectors. I observe no statistically significant effect for the construction sector.

TABLE 3.3:
Effects on individuals primary sector

	(1) Agriculture	(2) Manufacturing	(3) Construction	(4) Retail	(5) Transport	(6) Educ and public services
Panel A : IV results						
Rural elec. rate	-1.331***	0.244***	-0.057	0.216***	0.094***	0.670***
	(0.216)	(0.080)	(0.052)	(0.066)	(0.027)	(0.095)
Rural elec. rate * Tariff	2.017***	-0.693**	0.282	-0.471**	-0.177*	-1.288***
	(0.652)	(0.284)	(0.184)	(0.199)	(0.106)	(0.277)
Tariff	-1.537*	1.179***	-0.255	0.239	0.027	1.088***
	(0.859)	(0.421)	(0.242)	(0.235)	(0.124)	(0.363)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Panel B : OLS results						
Rural elec. rate	-0.040	0.007	-0.016*	0.033***	0.005	0.029**
	(0.029)	(0.013)	(0.009)	(0.009)	(0.004)	(0.013)
Rural elec. rate * Tariff	-0.273***	0.110***	0.070***	-0.005	0.008	-0.032
	(0.054)	(0.027)	(0.022)	(0.016)	(0.010)	(0.022)
Tariff	0.149	-0.042	-0.003	0.061**	-0.010	0.044
	(0.094)	(0.043)	(0.031)	(0.027)	(0.015)	(0.044)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Observations	445421	445421	445421	445421	445421	445421

Notes : Standard errors in parentheses are clustered at the village level. District level controls include the average rainfall and temperature, banks available per 1000 persons and the poverty gap as a measure of the level of poverty of districts. Dependant variables take value 1 if the individual works in the identified sector and 0 otherwise. All regressions include district and year fixed-effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

These findings suggest that when the domestic economy tends to be open to international trade, with a decline of the level of tariffs applied to foreign products, an increase of its agricultural productivity may trigger a labor reallocation from the agricultural sector towards other sectors of the economy. Hence, it appears that a tariff reduction may have an income effect that allows agricultural households to invest in the adoption of electricity for irrigation. [48] In fact, the reduction of tariffs may impact the prices of consumption and intermediate input goods ; allowing farmers to increase their saving and to improve their technology used for agricultural activities. This technology adoption may have released household labor from farming activities, allowing them to operate in other sectors of the economy.

Combining these results, Table 3.3 also suggest opposing evidence to Ricardo's theory of comparative advantage. In fact, in his theory, Ricardo suggests that a small open economy should specialize in the sector where it has a comparative advantage. While there is a labor reallocation towards other sectors in a situation of closed economy. The increase of the agricultural productivity generated by electrification in India has the potential to increase the comparative advantage of the agricultural sector. However, my results suggest that with the decrease of tariff level and an openness to international trade, individuals may have taken advantage of the agricultural productivity increase to diversify their economic activities, instead of specializing in the agricultural sector. In the remainder of the analysis, I show additional results that confirm the individual level results.

Households primary sector

I estimate the same regression as in the individual level analysis and see how households industry of activity is affected following a tariff decrease and an increase in the use of electri-

48. This explanation is consistent with Fajgelbaum and Khandelwal [2016] who show that trade openness has a greater positive effect on poor consumers through changes in their cost of living and an increase of their real income.

city. Results are presented in Table 3.4 and I observe the same pattern as in the individual level analysis. More specifically, I note a decrease of the number of households that have their activity in the agricultural sector. The point estimate in column (1) suggests that a one percentage point increase in electrification implies a decrease of households operating in the agricultural sector by 2.36 p.p, for districts in which tariff decrease more. For districts that experience the average or lowest decline of tariff, I notice a decrease by 2.07 p.p and 1.78 p.p respectively. Column (2) shows the opposite for the manufacturing, with an increase of households that have their activity in that sector by 0.51 p.p, 0.43 p.p and 0.36 p.p for districts at the highest, average and lowest decline of tariff. I also notice an increasing share of households operating in the transport and education & public services sectors.

TABLE 3.4:
Effects on households primary sector

	(1) Agriculture	(2) Manufacturing	(3) Construction	(4) Retail	(5) Transport	(6) Educ and public services
Panel A : IV results						
Rural elec. rate	-1.131***	0.198***	-0.087*	0.131**	0.099***	0.652***
	(0.194)	(0.064)	(0.051)	(0.065)	(0.036)	(0.105)
Rural elec. rate * Tariff	2.394***	-0.606***	0.142	-0.272	-0.415***	-1.118***
	(0.570)	(0.234)	(0.191)	(0.203)	(0.135)	(0.284)
Tariff	-2.212***	0.812**	0.091	0.220	0.207	1.088**
	(0.783)	(0.368)	(0.266)	(0.246)	(0.162)	(0.447)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Panel B : OLS results						
Rural elec. rate	-0.052	0.014	-0.017	0.039***	0.003	0.005
	(0.033)	(0.012)	(0.010)	(0.011)	(0.006)	(0.021)
Rural elec. rate * Tariff	-0.091	0.061**	0.007	0.006	-0.011	-0.031
	(0.061)	(0.027)	(0.025)	(0.022)	(0.015)	(0.032)
Tariff	0.003	-0.055	0.046	0.036	-0.009	0.104
	(0.105)	(0.038)	(0.035)	(0.034)	(0.021)	(0.069)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Observations	97923	97923	97923	97923	97923	97923

Notes : Standard errors in parentheses are clustered at the village level. District level controls include the average rainfall and temperature, banks available per 1000 persons and the poverty gap as a measure of the level of poverty of districts. Dependant variables take value 1 if the household has its primary activity in the identified sector and 0 otherwise. All regressions include district and year fixed-effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Number of days worked in each sector per week

The increase of electrification may have allowed individuals/households to reallocate their activities from the agricultural sector to non-agricultural sectors. Another possibility for them is to reduce the number of days worked per week in farming activities and to invest their remaining work time in non-agricultural activities. I investigate this issue by estimating equation (3.2) with the number of days worked in each sector per week, as the dependant variable. For instance, in column (1) of Table 3.5, the dependant variable is the number of days worked in the agricultural sector. I find no significant effect of a decrease of the number of days worked in that sector. However, I notice an increase for the retail and education & public services sectors, following an increase of electrification and a reduction of tariff. Coefficients estimated in column (6) are the most precisely estimated and are significant at 1% level. They suggest that a one percentage point increase in rural electrification implies an increase of the number of days worked in the education & public services sector by *2.409 - 4.866 ** Δ*Tariff*. To give a sense of the marginal effect, a one percentage point increase in rural electrification increase the number of days worked in the education & public services sector by 4.91, 4.31 and 3.73 days per week for districts experiencing the largest, the average and the lowest decrease of tariff respectively. Regarding the retail sector, I notice an increase of the number of days worked per week by 1.69, 1.46 and 1.24. These results are consistent with Emerick [2018] and McMillan and Harttgen [2014] who show that an increase of agricultural productivity may allow labor reallocation from the agricultural sector to the local service sector.

Table 3.5:
Effects on the number of days worked per week

	(1) Agriculture	(2) Manufacturing	(3) Construction	(4) Retail	(5) Transport	(6) Educ and public services
Panel A : IV results						
Rural elec. rate	-0.801	0.717***	0.047	0.730***	0.195**	2.409***
	(1.312)	(0.262)	(0.178)	(0.234)	(0.095)	(0.350)
Rural elec. rate * Tariff	-1.214	-1.062	0.726	-1.864**	-0.445	-4.866***
	(3.728)	(0.895)	(0.615)	(0.732)	(0.364)	(0.937)
Tariff	2.845	1.841	-0.816	1.075	0.076	4.772***
	(5.121)	(1.130)	(0.767)	(0.887)	(0.441)	(1.289)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Panel B : OLS results						
Rural elec. rate	0.099	0.004	-0.022	0.113***	0.011	0.072
	(0.214)	(0.046)	(0.025)	(0.032)	(0.015)	(0.066)
Rural elec. rate * Tariff	-0.695*	0.255***	0.148**	-0.017	0.018	-0.039
	(0.416)	(0.097)	(0.062)	(0.066)	(0.034)	(0.090)
Tariff	0.384	-0.223	0.020	0.138	0.021	-0.095
	(0.665)	(0.155)	(0.089)	(0.102)	(0.050)	(0.237)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Observations	445421	445421	445421	445421	445421	445421

Notes : Standard errors in parentheses are clustered at the village level. District level controls include the average rainfall and temperature, banks available per 1000 persons and the poverty gap as a measure of the level of poverty of districts. Dependant variables represent the number of days worked in the identified sector. All regressions include district and year fixed-effects.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

10.2 Robustness checks

Alternative measure of tariff

Recall that the tariff of non-traded goods are put to zero in Topalova [2010]'s database. For this reason, the tariff measure that I use above is sensitive to the share of non-traded goods in the consumption basket of households. In this section I present results of an alternative measure of tariff : the interaction between the national ad-valorem tariff and the consumption share of traded goods. [49] The mean of the alternative measure is around 0.52 in 1988 and declined up to 0.14 in 2000. The regression results are presented in the annexe section in Table 3.6 to Table 3.8. Quantitatively, the magnitude of the coefficients in Table 3.6 are similar to the ones presented in Table 3.3, showing the effect on individuals primary sector. Qualitatively, the coefficients on the tariff measure are not significant and suggest that the districts' exposure to tariff computed with only traded goods does not affect the labor share of sectors. But the effect of electrification conditional on the tariff level is captured by the coefficient of the rural electrification rate and of the interaction term. These coefficients are both statistically significant for the regression results presented in column (1), (4), (5) and (6). For instance, the estimated coefficient in column (1) for the rural electrification rate variable is -1.475 against -1.331 in the baseline (column 1 of Table 3.3). For the interaction term it is 1.736 against 2.017. These numbers correspond to a decrease of the agricultural labor share by 2.54 p.p, 2.14 p.p and 1.9 p.p for districts experiencing the highest, the average and the lowest decline of the alternative measure of tariff. Regarding the increase of the retail, transport and education & public service sectors' labor share, results remain similar to the baseline presented in Table 3.3.

49. The difference between this measure of tariff and the one used in the baseline results is that the consumption share of non-traded goods are not considered in the computation of this alternative measure. Only traded goods that compose the consumption basket of households are considered. Thus, this alternative measure is not mechanically influenced by the size of the consumption share of non-traded goods.

Table 3.7 presents results on households' industry of activity. They are similar to the baseline presented in Table 3.4. The same conclusion holds for the number of days worked by individuals in each sector (Table 3.8).

Exclusion restriction of the Hydro instrument

The instrumental variable estimations presented above are valid if and only if the hydroelectric supply shifts affect labor reallocation between sectors only through a variation in the rural electrification rate. To test this hypothesis, I restrict my sample to individuals living in districts where the variation of the rural electrification rate remain constant between 1988 and 2000. I consider that the electrification rate is constant for districts that have a variation of less than 10 percentage point.[50] I have identified 50 districts that are in this situation. I, then, perform a reduced form regression ; i.e the labor share of each sector on the hydro instrument, using this restricted sample. In this case, the exclusion restriction is verified if the coefficients on the hydro instrument for each specification are not statistically significant. Table 3.9 in the annexe section shows that the district and state level instruments are not statistically significant for all regressions, except in column (2) where the state level instrument is significant at 10% level. These results suggest that the hydroelectricity supply shifts may affect the labor share of sectors through a variation of the electrification rate of districts. I am, therefore, confident in claiming that the exclusion restriction hypothesis is not violated.

50. The average variation of the rural electrification rate between 1988 and 2000 is 25 percentage point (Table 3.1).

Falsification tests

I reshuffle the measure of rural electrification over the entire sample 1000 times and perform 1000 regressions of equation (3.2). I, separately, do the same operation with the measure of tariff. The objective of the exercise is to show that my baseline results do not capture a random correlation between the outcome variables and the variables of interest. I report in Table 3.10 in the annexe section the share of regressions that produce statistically significant estimates at 1%, 5% and 10% level. Of the 1000 regressions performed, only 0.1% produced significant estimates at 1% level for my three variables of interest (Column 1). The percentage increase for each variable at 5% and 10% significance level in column (2) and (3) respectively. Column (4) to (6) show similar results when I reshuffle the tariff measure. Overall, Table 3.10 confirms that the combined effect of the rural electrification rate and the tariff measure of districts are determinant in the reallocation of labor between sectors.

11 CONCLUSION AND DISCUSSION

This paper is interested on the empirical evidence that trade policy affects the relationship between agricultural productivity and structural transformation. It has shown in the case of India that an increase of the electricity use for irrigation combined with a decrease of tariff level have a statistically significant effect on labor reallocation in rural areas from the agricultural sector towards the non-agricultural sectors. More specifically, the combine effect of tariff reduction and agricultural productivity cause an increase of the labor share in the manufacturing, retail, transport and education and public services sectors. Additional results on households' primary sector of activity confirm the individual level results. Although there is no statistically significant on a decrease of the number of days worked in the the agricultural sector, the sign on the coefficient suggests that individuals work less in this sector. And I notice a statistically significant increase of the number of days worked in the retail and education and public services sectors.

My results are an argument in favor of globalization and agricultural technology development in developing countries because they show that their combine effect can affect positively the development of non-agricultural sectors. And they show opposite evidence to the theory of comparative advantage that states that in open economies, a comparative advantage in the agricultural sector can dampen industrial expansion.

It is, however, important to note that with the instrumental variable approach that I use, I estimate a Local Average Treatment Effect (LATE). In other words, I capture the effect on individuals/households that use electricity for irrigation because of the hydroelectricity supply shift. That is why, one must be aware of this fact when generalizing this results at the national level or when adopting such policies in other developing countries.

12 APPENDIX : TABLES

Table 3.6:
Robustness check : Effect on individuals primary sector

	(1) Agriculture	(2) Manufacturing	(3) Construction	(4) Retail	(5) Transport	(6) Educ and public services
Panel A : IV results						
Rural elec. rate	-1.475***	0.323***	0.00805	0.206***	0.101***	0.622***
	(0.260)	(0.116)	(0.0647)	(0.0770)	(0.0355)	(0.107)
Rural elec. rate * Tariff	1.736***	-0.327	0.349**	-0.498***	-0.213**	-0.946***
	(0.605)	(0.252)	(0.165)	(0.179)	(0.0923)	(0.241)
Tariff	-1.102	0.246	-0.425	0.351	0.150	0.364
	(0.989)	(0.438)	(0.269)	(0.265)	(0.134)	(0.337)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Panel B : OLS results						
Rural elec. rate	-0.0614**	0.00905	-0.0165**	0.0283***	0.00773*	0.0233*
	(0.0280)	(0.0126)	(0.00840)	(0.00814)	(0.00399)	(0.0128)
Rural elec. rate * Tariff	-0.271***	0.122***	0.0930***	0.0205	0.00306	-0.0121
	(0.0478)	(0.0259)	(0.0192)	(0.0155)	(0.00858)	(0.0157)
Tariff	0.00877	-0.121***	-0.0608**	0.0234	0.0270*	0.0145
	(0.0880)	(0.0435)	(0.0286)	(0.0279)	(0.0147)	(0.0345)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Observations	443299	443299	443299	443299	443299	443299

Notes : This table is equivalent to Table 3.3. I use an alternative measure of tariff. Standard errors in parentheses are clustered at the village level. District level controls include the average rainfall and temperature, banks available per 1000 persons and the poverty gap as a measure of the level of poverty of districts. Dependant variables take value 1 if the individual works in the identified sector and 0 otherwise. All regressions include district and year fixed-effects.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7:
Robustness Check : Effect on households primary sector

	(1) Agriculture	(2) Manufacturing	(3) Construction	(4) Retail	(5) Transport	(6) Educ and public services
Panel A : IV results						
Rural elec. rate	-1.005***	0.212**	-0.0754	0.0939	0.0953**	0.508***
	(0.199)	(0.0837)	(0.0618)	(0.0722)	(0.0403)	(0.0972)
Rural elec. rate * Tariff	2.266***	-0.466**	0.0262	-0.378*	-0.393***	-0.853***
	(0.566)	(0.220)	(0.182)	(0.204)	(0.127)	(0.270)
Tariff	-2.000**	0.400	0.297	0.449	0.175	0.497
	(0.949)	(0.381)	(0.295)	(0.310)	(0.198)	(0.392)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Panel B : OLS results						
Rural elec. rate	-0.0718**	0.0174	-0.0248**	0.0365***	0.00364	0.0177
	(0.0297)	(0.0121)	(0.00996)	(0.0101)	(0.00594)	(0.0169)
Rural elec. rate * Tariff	-0.152***	0.0673***	0.0292	0.0248	-0.0160	0.0304
	(0.0531)	(0.0255)	(0.0224)	(0.0206)	(0.0135)	(0.0231)
Tariff	0.0270	-0.0626	-0.0524	0.0311	0.0325	0.0105
	(0.0917)	(0.0436)	(0.0350)	(0.0364)	(0.0225)	(0.0375)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Observations	97424	97424	97424	97424	97424	97424

Notes : This table is equivalent to Table 3.4. I use an alternative measure of tariff. Standard errors in parentheses are clustered at the village level. District level controls include the average rainfall and temperature, banks available per 1000 persons and the poverty gap as a measure of the level of poverty of districts. Dependant variables take value 1 if the individual works in the identified sector and 0 otherwise. All regressions include district and year fixed-effects.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8:
Robustness check : Effect on number of days worked per week

	(1) Agriculture	(2) Manufacturing	(3) Construction	(4) Retail	(5) Transport	(6) Educ and public services
Panel A : IV results						
Rural elec. rate	-2.021	1.089***	0.159	0.827***	0.187	2.076***
	(1.477)	(0.370)	(0.198)	(0.279)	(0.121)	(0.373)
Rural elec. rate * Tariff	0.579	-0.184	0.815	-1.422**	-0.649**	-2.986***
	(3.458)	(0.903)	(0.514)	(0.659)	(0.308)	(0.788)
Tariff	-1.836	-0.358	-1.033	0.252	0.632	0.469
	(6.115)	(1.541)	(0.762)	(0.985)	(0.458)	(1.141)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Panel B : OLS results						
Rural elec. rate	0.0633	0.0166	-0.0357	0.109***	0.0152	0.0744
	(0.204)	(0.0430)	(0.0237)	(0.0298)	(0.0143)	(0.0548)
Rural elec. rate * Tariff	-0.864**	0.247***	0.236***	0.0213	0.0146	-0.0775
	(0.385)	(0.0935)	(0.0554)	(0.0607)	(0.0311)	(0.0512)
Tariff	-0.264	-0.291*	-0.256***	0.173	0.104*	0.126
	(0.615)	(0.170)	(0.0881)	(0.107)	(0.0554)	(0.0985)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Observations	443299	443299	443299	443299	443299	443299

Notes : This table is equivalent to Table 3.5. I use an alternative measure of tariff. Standard errors in parentheses are clustered at the village level. District level controls include the average rainfall and temperature, banks available per 1000 persons and the poverty gap as a measure of the level of poverty of districts. Dependant variables represent the number of days worked in the identified sector. All regressions include district and year fixed-effects.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 3.9:
Placebo : Exclusion restriction of the hydro instrument

	(1) Agriculture	(2) Manufacturing	(3) Construction	(4) Retail	(5) Transport	(6) Educ and public services
$Hydro_{d,t}$ ♠	-1.878	0.730	-0.117	1.112	0.305	1.126
	(1.988)	(0.941)	(0.570)	(0.709)	(0.408)	(0.713)
$Hydro_{s,t}$ ♣	-0.331	0.507*	-0.245	-0.019	-0.127	-0.320
	(0.587)	(0.304)	(0.172)	(0.251)	(0.161)	(0.240)
$Hydro_{d,t} * PredictedTariff$	4.059	-1.858	0.244	-2.438	-0.708	-2.570
	(4.430)	(2.084)	(1.282)	(1.580)	(0.913)	(1.601)
Predicted Tariff	2.592	1.661	1.337*	-1.684	-0.394	-1.185
	(3.482)	(1.603)	(0.705)	(1.230)	(0.599)	(1.141)
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year F-E	Yes	Yes	Yes	Yes	Yes	Yes
Observations	68734	68734	68734	68734	68734	68734
Mean of dep. var	0.742	0.058	0.023	0.05	0.013	0.054
	(0.437)	(0.233)	(0.152)	(0.217)	(0.112)	(0.227)

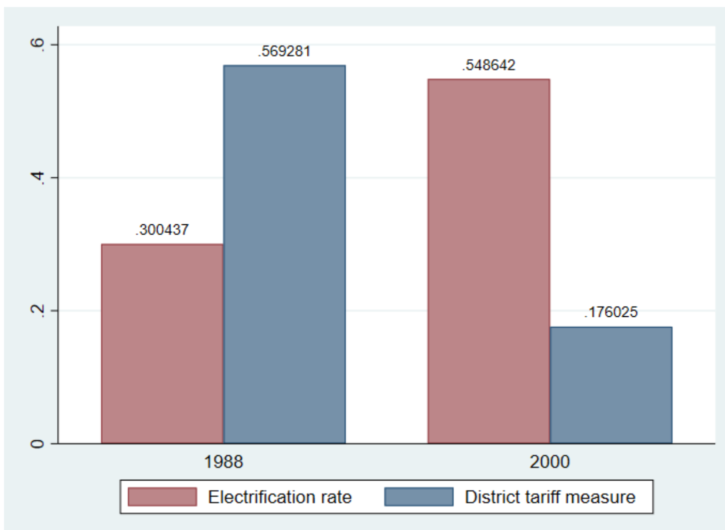
Notes : Standard errors in parentheses are clustered at the village level. District level controls include the average rainfall and temperature, banks available per 1000 persons and the poverty gap as a measure of the level of poverty of districts. Dependant variables take value 1 if the individual works in the identified sector and 0 otherwise. All regressions include district and year fixed-effects. Predicted tariff represents the predict of the regression of the tariff measure on geographical accessibility variables and is the instrument of the tariff measure. ♠ is the district level instrument for the Rural elec. rate : $Hydro_{s,t} * Water_{d,s,1970}$. ♣ is Allcott et al. [2016a] instrument at the state level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.10:
Falsification Tests with Agri. labor share as dependant variable

Placebo	Reshuffle Rural electrification rate: Share of estimations with statistical significance at : 1% (1)	5% (2)	10% (3)	Reshuffle Tariff measure: Share of estimations with statistical significance at : 1% (4)	5% (5)	10% (6)
Rural elec. rate	0.001	0.023	0.066	0	0.012	0.044
Rural elec. rate * Tariff	0.001	0.019	0.058	0	0.014	0.037
Tariff	0.001	0.02	0.058	0	0.021	0.047
District level controls	Yes	Yes	Yes	Yes	Yes	Yes
District and year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	445421	445421	445421	445421	445421	445421

Notes : The table shows the share of statistically significant results over 1000 randomizations, where the rural electrification rate and the measure of tariff are randomized, separately, over the entire sample. In each regression standard errors are clustered at the village level and include district level controls such as the average rainfall and temperature, banks available per 1000 persons and the poverty gap as a measure of the level of poverty of districts. All regressions include district and year fixed-effects. The dependant variable takes value 1 if the individual works in the agricultural sector and 0 otherwise.

13 APPENDIX : FIGURES



Note : The maroon bars show the evolution of the rural electrification rate while the blue ones show the decrease of the districts' tariff measure between 1988 and 2000.

FIGURE 3.8 Variation of the districts' tariff measure and the rural electrification rate between 1988 and 2000

CHAPITRE 4
THE LONG-TERM EFFECTS OF UNEXPECTED INTERRUPTIONS IN COMPULSORY SCHOOLING

1 AVANT-PROPOS

Ce chapitre a été co-écrit avec A. Bernabe, M. Pelli et J. Tschopp et a déjà été soumis pour publication à une revue scientifique. J'ai eu l'idée pour cet article en discutant du sujet avec mon superviseur, qui m'a ensuite fourni les données pour la mettre en place. Mon rôle, après avoir eu l'idée, a été de mettre en place les données, effectuer une analyse préliminaire et produire un premier draft du chapitre sous la supervision des Pr. Pelli et Tschopp. Ce draft a ensuite été retravaillé par Pr. Pelli et Pr. Tschopp pour en améliorer la qualité. Ils ont également ajouté des tests de robustesse. Angélique Bernabé a travaillé avec moi au traitement de la base de données, et à une partie de l'analyse préliminaire.

Citation :

Bernabé, Angélique and Diop, Boubacar and Pelli, Martino and Tschopp, Jeanne, The Long-Term Effects of Unexpected Interruptions in Compulsory Schooling (September 6, 2021). Available at SSRN : `https://ssrn.com/abstract=3806538` or `http://dx.doi.org/10.2139/ssrn.3806538`

2 RÉSUMÉ

Ce papier étudie les impacts à long terme des interruptions inattendues de la scolarité obligatoire. En utilisant les tempêtes comme choc exogène, nous examinons comment les interruptions de la scolarité obligatoire affectent les résultats scolaires et le type d'activité exercée par les individus lorsqu'ils deviennent de jeunes adultes. Nous construisons une mesure continue unique de l'exposition aux tempêtes durant l'enfance qui varie selon la cohorte de l'année de naissance et le district pour les jeunes adultes résidant dans les zones rurale et urbaine en Inde. Nous constatons que les tempêtes ont un impact perturbateur important sur l'éducation. Dans les districts exposés aux vents les plus puissants, les estimations impliquent que les enfants ont 9% plus de chances d'accumuler un retard scolaire et 6,5% moins de chances d'obtenir un niveau d'éducation supérieur (au-delà du secondaire). À long terme, ces retards ont un impact sur le type d'activité sur le marché du travail que ces individus exercent. En utilisant l'exposition de l'enfance aux tempêtes comme variable instrumentale, nous constatons qu'un retard d'un an dans l'éducation entraîne une baisse de 42,6% de la probabilité d'accéder à des emplois salariés réguliers. Nous déterminons que l'impact des tempêtes sur l'éducation passe par un choc de revenu négatif permanent.

3 ABSTRACT

In this paper we quantify the long-run impacts of childhood exposure to storms on education and labor market activities in urban and rural India. The identification strategy relies on an original continuous measure of exposure to storms during compulsory schooling that varies by birth-year cohort and district. Our results suggest that storms have substantial disruptive impacts on education and labor market prospects, causing a deskilling of the affected regions. The estimates are particularly sizable in the case of severe cyclonic storms which, because of climate change, have surged over the past few years. In this case, our findings imply an increase of 18 percentage points of the probability of accumulating an educational delay of at last one year, a 4.8 percentage points increase in the probability of no formal schooling and a fall of 8.1 percentage points in the probability of completing post-secondary education. In the long run, childhood exposure to storms has an impact on the type of labor market activity performed, causing a reduction in the probability of accessing regular salaried jobs while increasing the odds of performing domestic duties as primary activity.

Keywords : schooling interruptions, education, labor markets, storms

JEL Codes : I25, Q54.

4 INTRODUCTION

Over the last 10 years, weather-related natural catastrophes generated on average over 200 billion US dollars of damages per year globally. The main culprits of these damages are tropical cyclones. [51] Emanuel [2021] shows that an increase in greenhouse gases in the atmosphere is directly responsible for an increase in frequency and intensity of tropical cyclones. Evidence of this link and growing concerns about climate change imply that the cost of these events is likely to increase dramatically over the coming years and call for a thorough understanding of their impacts on the economy.

The short- and long-term effects of natural disasters on economic growth have been studied extensively [e.g. Cavallo, Galiani, Noy, and Pantano, 2013, Cavallo and Noy, 2010, Dell, Jones, and Olken, 2014, Strobl, 2011]. While there is a general consensus on the negative contemporaneous consequences of natural disasters, recent findings are in disagreement regarding their long-term effects. [52] The majority of studies relates the path of GDP growth to physical capital reconstruction and potential technological upgrading, yet, to the best of our knowledge, causal evidence of the long-run effects of the impact of natural disasters on human capital formation is scant.

51. `https://www.swissre.com/media/news-releases/nr-20201215-sigma-full-year-2020-preliminary-natcat-loss-estimates.html`

52. Hsiang and Jina [2014] summarize the literature that describes the long-term evolution of GDP per capita in the aftermath of a natural disaster. The authors put foward four hypotheses : *(i) creative destruction*, *(ii) build back better*, *(iii) recovery to trend* and, *(iv) no recovery*. In the long run, each of these hypotheses predicts a different level of GDP per capita.

The literature has long established that education is an important determinant of an individual's earnings and that, in the aggregate, human capital contributes to the economic growth of a nation. [53] As a consequence, if natural disasters hinder academic achievements, we may still expect economic growth to slow down in the long run, even if environmental disruption stimulates innovation and assets are replaced with newer and more productive vintages.

In this paper we quantify the long-run impacts of childhood exposure to storms over the course of compulsory schooling on both educational attainments and the type of activity performed by individuals in young adulthood in urban and rural India. We focus on India which is one of the most impacted regions in the world, with over 370 million people affected yearly (roughly one in four people). [54] Natural disasters can disrupt education through two channels : the supply and the demand for schooling. The former channel likely operates through the destruction of schools and road infrastructures, which have been shown to play a key role in promoting education [see for instance Duflo, 2001, Jaume and Willèn, 2019]. The demand channel may be linked to the impact of storms on the psychological health of children [see Kar and Bastia, 2006, Neria, Nandi, and Galea, 2008] and/or households' income.

To study the long-run impacts of childhood exposure to storms, data requirements are high. For each individual we need information on both, current outcomes and exposure to storms during compulsory schooling. We combine data from the 2018 release of the Periodic Labour Force Survey (PLFS) with storms' best tracks data from the National Oceanic and Atmospheric Administration (NOAA) over the period 1990-2010. Our identification strategy relies on an original measure of childhood exposure to storms constructed from exogenous variations in wind exposure

53. See for instance Topel [1999] for a study on the role of human capital in economic growth. See Card [2001] for a survey of papers that attempt to identify the impact of education on labor market earnings using supply-side features of the education system (e.g. compulsory schooling laws or differences in the accessibility of schools) as determinants of schooling outcomes. For more recent evidence see Chetty, Friedman, Hilger, Saez, Whitmore Schanzenbach, and Yagan [2011].

54. See `https://ncrmp.gov.in/cyclones-their-impact-in-india/`

across birth-year cohorts and districts during compulsory schooling. [55] We proceed in two steps. First, for each year between 1990 and 2010, we build an index of yearly district exposure to storms that accounts for the force exerted by winds on physical structures at the district's geographical centroid. Second, for each district and birth-year cohort aged between 23 and 33 in 2018, we aggregate the index over years of compulsory schooling and obtain a continuous treatment that varies by birth-year cohort and district.

We find that exposure to storms over the course of compulsory schooling impacts long-term educational attainments and the type of activity performed in young adulthood. In the case of severe tropical storms, the estimates imply an average schooling delay of close to three months and an 18 percentage points increase in the probability of repeating a year or dropping out of school. While affected children still complete primary schooling, the delays accumulated over time translate in a lower propensity to complete higher education. Estimates suggest that children exposed to storms face a decrease of 8.1 percentage points in the probability of completing post-secondary education. Concurrently, the probability of having no formal schooling increases by 4.8 percentage points. These delays also manifest themselves later on, once affected individuals reach young adulthood. Positive exposure reduces the likelihood of working a formal job or being self-employed, while it increases the probability of performing domestic duties as primary activity. Hence, our results suggest that tropical storms can lead to a deskilling of future generations and, in consequence, a widening of the skill and wage gaps across affected and sheltered regions. If, on the one hand, storms foster innovations and technological advancement, on the other hand, our findings indicate that they have negative effects on human capital formation, therefore potentially compromising economic growth in the long run.

55. The occurrence of storms is random and unpredictable [see for instance Elsner and Bossak, 2001, Pielke, Landsea, Mayfield, Laver, and Pasch, 2008], and the literature has shown that individuals do not account for it in their decision process [Dessaint and Matray, 2017, Elsner and Bossak, 2001, Pielke et al., 2008, Wu, Lindell, and Prater, 2014]. We provide a detailed discussion regarding the exogeneity of storms to economic activity in Pelli and Tschopp [2017] and Pelli, Tschopp, Bezmaternykh, and Eklou [2020].

Our paper is closely related to Maccini and Yang [2009] who adopt a similar methodology to quantify the impact of early-life (0 to 5 years) rainfall shocks on adult outcomes (health, schooling and wealth) in Indonesia. While our identification strategy is similar in essence, we use a different type of shock, focusing on exposure to storms during years of compulsory schooling. Our paper informs on public policies that respond to extreme events whereas their findings arise from more typical year-to-year variation in rural households' economic conditions (variation in agricultural output) around the time of birth. The authors find evidence that early-life rainfall is an important determinant of adult socioeconomic status.

The literature on the effect of natural disasters on education has mainly focused on the contemporaneous effects of disasters. An exception is Caruso [2017] which, in line with our results, finds that children suffer long-lasting negative effects from natural disasters. However, the paper examines all types of natural disasters occurring in Latin America over the last 100 years and identifies them with dummy variables. Deuchert and Felfe [2015] look at a super typhoon on Cebu island in the Philippines and show a negative effect on children's education, probably due to a shift in households' spending towards reconstruction. Groppo and Kraehnert [2017] look at the impact of severe winters in Mongolia and find that children's education suffers, likely because severe winters act as a negative income shock. Rosales-Rueda [2018] shows lower test score results for children affected by floods while in utero in Ecuador. Spencer, Polachek, and Strobl [2016] look at the contemporaneous effect of cyclones on educational results in the Caribbeans, which turn out to be negative.[56]

56. Many papers focus on developed countries. Karbownik and Wray [2019] investigates the impact of exposure to cyclones on fetal and early life in the US and find a negative income effect in adulthood for white males. Billings, Gallagher, and Ricketts [2020] show a decrease in enrolment numbers and in graduation rates and, Sacerdote [2012] finds an initial decrease in test scores of students affected by Hurricanes Katrina and Rita but a subsequent increase for students moving out of Louisiana to states with better school systems. Groen, Kutzbach, and Polivka [2020] study the effects of the same hurricanes on employment and earnings. While job losses generate short-term fall in earnings, they find that, in the long run, affected regions experience gains which are due to both a contraction of labor supply and an increase in labor demand.

Finally, this paper also speaks to the literature on education in developing countries. Evidence suggests that improving school attendance, subsidizing textbooks [Glewwe, Kremer, and Moulin, 2009] or even increasing the number of teachers [Banerjee, Suraj, and Kremer, 2004] does not necessarily ameliorate learning. Therefore, recent findings suggest that improving the quality of teaching is a first-order concern [see for instance Banerjee, Cole, Duflo, and Linden, 2007] and that policies promoting school enrolment should, at the very least, be coupled with interventions improving the pedagogy or curriculum of schooling. [57] While we do not provide direct evidence on the quality of teaching, our results seem to indicate that school attendance remains very important as the delays which appear to result from absenteeism or school removals do have long-term consequences on the probability of completing higher education and on future labor market outcomes.

5 DATA

Our empirical analysis uses two sources of data : *i)* the 2018 release of the PLFS – used to measure educational delays and labor market variables, and *ii)* tropical storms data from the NOAA – used to construct an index of childhood exposure to storms.

5.1 Individual and household data

The PLFS is an individual- and household-level representative survey of the Indian population collected by the National Sample Survey Office (NSSO) of the Ministry of Statistics and Program Implementation. The survey provides a variety of information on individuals' characteristics such as age, gender, educational level and the number of years spent at school. In India, children typically start school at the age of 6. Without delays, compulsory schooling lasts 9 years, i.e. until a child is 15 years old. Table 4.1 summarizes the schooling system, including the various paths to higher education. [58] Column (1) indicates the number of years needed to complete each

57. Duflo [2001] and Duflo [2004] estimate the impact of school construction on education and labor market outcomes in Indonesia. For more references, see Glewwe and Kremer [2006]. See Chetty et al. [2011] for a study on the long-run impacts of early childhood education in the U.S.

58. More details on the educational system of India and how it compares to other systems can be found here : `https://wenr.wes.org/2018/09/education-in-india`

individual category of schooling. For graduate and postgraduate levels, the numbers correspond to the modal duration across disciplines. Column (2) shows the total cumulated number of years needed to complete any given level of education. For instance, middle school lasts 3 years. At the end of middle school, a child should have accumulated 8 years of education ; 5 years of primary and 3 years of middle school. The PLFS provides information on the highest level of education completed and on whether an individual earned a diploma/certificate. These two pieces of information allow us to infer the path of individuals who continued into higher education and compute the corresponding theoretical number of years of education (in the absence of an educational delay).

Table 4.1:
Schooling System in India

	Duration (1)	Cumulated Years of Education (2)
Lower education :		
Primary	5	5
Middle	3	8
Secondary	2	10
Higher secondary	2	12
Higher education :		
Path 1 :		
Diploma/certificate course	1	13
Path 2 :		
Graduate	3	15
Path 3 :		
Diploma/certificate course	1	13
Graduate	3	16
Path 4 :		
Graduate	3	15
Postgraduate and above	2	17
Path 5 :		
Diploma/certificate course	1	13
Graduate	3	16
Postgraduate and above	2	18

Notes : Column (1) shows the duration of each category of schooling. For *Graduate* and *Postgraduate*, the duration corresponds to the mode across disciplines. Column (2) gives the total number of years of education accumulated after completion of each category of schooling (and path in the case of higher education).

For each individual, we measure educational delay as the difference between the actual number of years in formal education and the minimum number of years needed in the schooling system to achieve the reported level of education. For example, suppose an individual reports seven years of formal schooling but has only completed primary school. This individual has a two-years educational delay, which may be caused either by repeating grades or by dropping out from a higher educational level (middle school, in this particular example).[59] Thus, our analysis will inform on whether storms increase educational delays but it will not be able to tell us anything about the likelihood of repeating grades versus dropping out of school. As an alternative measure of educational delay we also construct an indicator variable taking the value of 1 for individuals with positive educational delays and 0 otherwise.

The PLFS provides information on the primary activity status of individuals.[60] For instance, we know whether an individual's primary activity takes place in the formal labor market or at home (e.g. performing domestic duties – collecting vegetables, firewood, cattle feed, sewing, etc.). Included among formal labor market activities are regular work (i.e. work associated with a formal job and an employment contract), casual work (i.e. work with a daily or periodic contract only), self-employment, and unpaid family work (e.g. work in the family business/farm without pay). The survey also contains labor market indicators such as hours of work and earnings, yet this information only pertains to individuals who perform paid activities and report being part of the labor force.

Importantly, the PLFS provides information on the district of residence of individuals, which, combined with information on individuals' age, allows us to create a unique measure of childhood exposure to storms that vary by birth-year cohort and district. As we describe below, our measure is a continuous treatment taking into account the intensity of the storms to which children of a given cohort and living in a specific district were exposed over the course of compulsory schooling. Given the very small proportion of individuals migrating outside of their birth's district [see, for instance, Edmonds et al., 2010a, Gupta, 1987, Munshi and Rosenzweig, 2009, Topalova,

59. While it would be interesting to distinguish between both types of delays, the PLFS does not provide sufficient information to distinguish between the two cases.

60. Details and definitions can be found here (p.35) : `http://mospi.nic.in/sites/default/files/publication$_$reports/Annual$_$Report$_$PLFS$_$2018$_$19$_$HL.pdf`

2010], we assume that individuals completed their compulsory schooling in the same district in which they are living in 2018. [61] This assumption is important for the construction of the childhood exposure index.

As benchmark age for young adulthood we choose the age of an individual at the time of completing postgraduate education (master degree). Without educational delays, obtaining a postgraduate degree takes 17 years. Children usually start school at the age of 6 and, therefore, young adulthood is reached at the age of 23. As a consequence, the youngest cohort considered in the paper was born in 1995 and should have completed compulsory schooling in 2010. The oldest cohort examined is dictated by the quality of satellite coverage. World Meteorological Organization (WMO)-sanctioned cyclone data for the North Indian Ocean only goes back to 1990. [62] As illustrated in Figure 4.1, this means that the oldest cohort we consider was born in 1985, and is 33 years old in 2018.

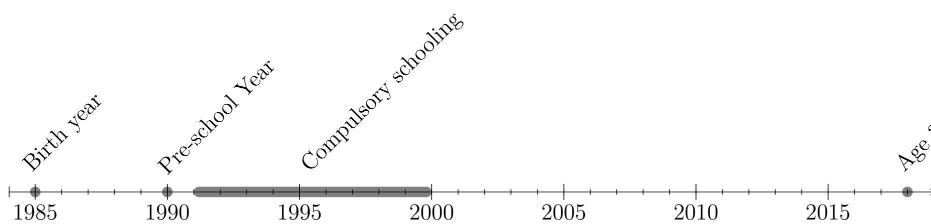


FIGURE 4.1 Oldest Cohort

Therefore, our analysis focuses on the 77,737 individuals born between 1985 and 1995 (i.e. the cohorts aged 23-33 in 2018) and storms which took place between 1990 and 2010.

61. Migration in India is low and, according to Topalova [2010], only less than 4% (13%) of rural (urban) individuals migrate out of district. We also compute our own migration numbers using the 64th round of the National Sample Survey (NSS) for the years 2007-2008. We find that only 3.5% of households (out of 125,578) have migrated within the last 365 days and that 1.3% have migrated permanently, out of which about half migrate out-of-district.

62. See `https://climatedataguide.ucar.edu/climate-data/ibtracs-tropical-cyclone-best-track-data`.

5.2 Childhood exposure to tropical storms

In order to understand how childhood exposure to storms impacts long-term education levels and labor market outcomes, we create an index based on storms' wind speed that varies by birth-year cohort and district. This measure captures storms occurring in the first nine years of compulsory schooling (starting at age six) and in the pre-school year. This additional year allows us to account for children born early in the year and, therefore, integrating school a year earlier. Childhood exposure to storms is computed as follows :

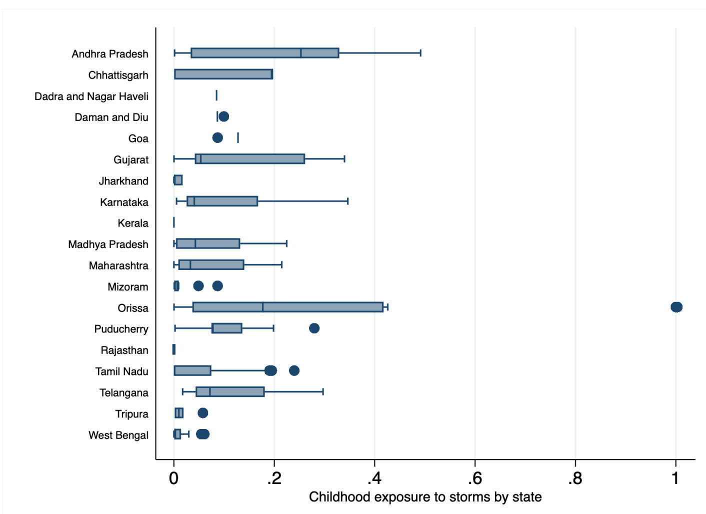
$$C_{bd} = \sum_{t=b+5}^{t=b+15} x_{dt}, \tag{4.1}$$

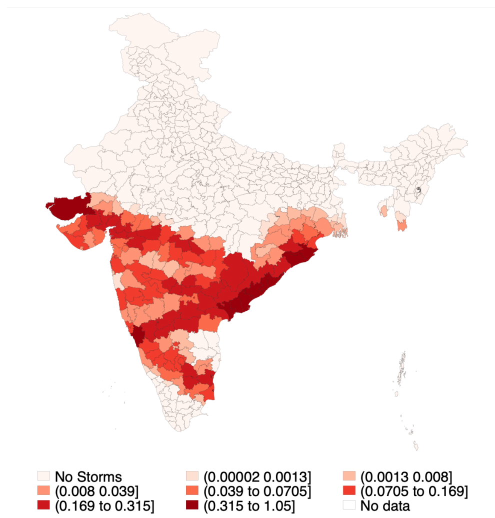
where b denotes a birth-year cohort, d a district, t a year. The variable x_{dt} is an index of yearly district exposure to storms and accounts for the force exerted by winds on physical structures. Details on the construction of x_{dt} are presented in Appendix B. Consider for instance the timeline of the oldest cohort (born in 1985). As illustrated in Figure 4.1, the index of childhood exposure, C_{bd}, sums district exposure to storms from 1990 (the pre-school year) up to 2000 (the end of the nine years of compulsory schooling) ; i.e. $C_{1985,d} = \sum_{t=1990}^{t=2000} x_{dt}$. Within birth-year cohort across district variation in the index results from the fact that at a given point in time, the exact same storm exerts different windspeed intensities at various locations, while some areas are sheltered. Accounting for wind speed lends us with a continuous treatment that varies across space, which is a considerable advantage in terms of identifying variation in state of relying on dummies or categorical treatments (e.g. a measure taking the value of one if an individual was exposed to a storm during the period of compulsory schooling). Within district across birth-year cohorts variation results from the fact that different cohorts may be subject to different storms over the course of compulsory schooling.

The left panel of Figure 4.2 presents the measure of childhood exposure to storms at the state level.[63] In our sample, children living in 28 out of the 35 Indian states experienced tropical storms between the ages of 5 and 15. Importantly, the boxplots show substantial variation in

63. Only states with positive exposures are included.

childhood exposure to storms within and across states, with Andhra Pradesh, Gujarat, Maharashtra, Orissa and Telangana displaying the largest median exposures. The right panel of Figure 4.2 provides a visualization of the distribution of C_{bd} across districts for the cohort born in 1987, with darker shades of red indicating higher exposures. The darkest shade indicates districts for which the index of childhood exposure falls above the 90th percentile in the distribution of C_{bd} for individuals born in 1987. Each shade contains 15% of the districts with a positive childhood exposure. The landlocked part of India in the North exhibits nearly zero exposure, which is consistent with storms' best track data which typically indicates a high concentration of storms along coastal areas. The map reveals that the cohort of 1987 living in the remainder of India experienced positive exposure to storms, with districts around the South-Eastern coast being the most affected.





Notes : The boxplots (left panel) describe the measure of childhood exposure to storms for positive values of exposure ($C_{bd} > 0$) and individuals born between 1985 and 1995 by state, listed in alphabetical order. The figure only shows states with positive exposure. The blue line in each box is the median. The lower bound of a box is the first quartile and the higher bound is the third quartile. The end of the left (right) whisker is the 1st percentile (99th percentile). Circles without a box mean that all observations are clustered around the median. The circles outside of the box represent outliers. The map (right panel) provides a visual illustration of childhood exposure to storms across districts for the cohort born in 1987. The darkest shades correspond to districts for which the index of childhood exposure falls above the 90th percentile in the distribution of C_{bd} in 1987. The other shades corresponding to positive exposures contain 15% of the districts each.

Figure 4.2 Childhood Exposure to Storms

In Table 4.2 we provide summary statistics for the main variables used in the paper. Panel A shows figures for the measure of childhood exposure to storms. 23,547 out of the 77,373 individuals included in the sample – roughly 30% – were exposed to storms over the course of compulsory schooling. In Panel B we present the list of individual controls that are used in the analysis. The sample is evenly split across genders, with a majority of Hindu households and 30% of first-born individuals.[64] Columns (1) and (2) of Panel C show the mean and standard deviation of dummy variables for the highest category of schooling completed by individuals. About 12% of our sample falls in the category *below primary*, 76% of these individuals did not attend (formal) school at all, while the remaining 24% are primary school dropouts. 9% completed primary school only, 22% middle school and 33% secondary school. Finally, 24% of the sample obtained a diploma (completed a certificate course) or obtained a post-/graduate degree. Columns (3) and (4) of Panel B show means of the variables with zero and positive exposure, respectively. The last column displays the difference between the latter two means and tests its statistical significance. Means do not statistically differ from each other for individuals falling in the categories *below primary* and *primary*. However, they differ in a statistically significant manner for individuals who completed higher levels of education.

Looking at Panel D, prima facie evidence suggests that, on average, the educational delay tends to be greater for individuals with positive exposure. Individuals with zero exposure experience an average delay of 0.47 years, which amounts to about 20 weeks.[65] The delay is on average two weeks longer for individuals with positive exposure (0.53 years, i.e. about 22 weeks). We check whether the same difference exists within educational categories, distinguishing between individuals who have completed at most primary, middle, secondary or higher educational levels. The difference in educational delays between the zero and positive exposure groups is particularly marked for individuals who did not go past middle school. For secondary and categories of schooling that fall into *higher education*, the sign of the relationship is opposite. With compulsory schooling lasting roughly until completion of secondary school, results in Panel C suggest that delays associated with storms tend to occur quite immediately, rather than appearing gradually over the years. Combining the results from Panel C and D, it appears that while the delays accumulated over the period of

64. The set of controls that we can use is highly restricted, because most household- and individual-level controls are likely to be affected by storms and would, therefore, be bad controls.

65. The school year lasts 42 weeks.

compulsory schooling do not prevent its completion, they have long-term impacts by reducing the likelihood of going past *middle* education levels. It is important to note that these are simple observations based on descriptive statistics, without accounting for different storm exposures and possible confounders.

Regarding labor market indicators, the descriptive statistics suggest that for the sub-sample of individuals who have positive salaries and report belonging to the labor force, positive exposure is associated with longer workweeks. Finally, the bottom of Panel D presents binary variables for the primary activity status of individuals. Among these individuals, the largest share, 35%, carries out domestic duties while only 18% has a formal job with a regular employment contract and salary.

Table 4.2:
Summary Statistics

	Mean	Std. Dev.	Min	Max	N	
	(1)	(2)	(3)	(4)	(5)	
Panel A : Exposure to storms						
C_{bd}	0.03	0.096	0	1.003	77,737	
$C_{bd} > 0$	0.099	0.154	1.23e-08	1.003	23,547	
Panel B : Controls						
Gender	0.49	0.50	0	1	77,737	
First born	0.29	0.45	0	1	77,737	
Hinduism	0.74	0.44	0	1	77,737	
	All		Zero exp.	Pos. exp.	Diff.	Diff.
	Mean	Std. Dev.	Mean	Mean	(3)-(4)	in weeks
	(1)	(2)	(3)	(4)	(5)	(6)
Panel C :						
Highest category of schooling completed (yes=1, no=0)						
Below primary♠	0.12	0.32	0.12	0.12	-0.001	
Primary	0.09	0.29	0.09	0.09	0.003	
Middle	0.22	0.41	0.22	0.21	0.01***	
Secondary	0.33	0.47	0.33	0.33	-0.004	
Above♣	0.24	0.43	0.24	0.25	-0.01***	
Observations	77,737		54,190	23,547		
Panel D : Main variables						
Educational delay						
Educational delay (yes=1, no=0)	0.30	0.46	0.29	0.33	-0.04***	
Educational delay (# of years)	0.48	0.88	0.47	0.53	-0.06***	2.52***
Primary	0.74	0.93	0.70	0.86	-0.16***	6.72***
Middle	0.59	0.67	0.53	0.75	-0.22***	9.24***
Secondary	0.30	0.71	0.31	0.26	0.05***	2.1***
Above	0.39	0.75	0.39	0.38	0.01	0.42
Observations	77,737		54,190	23,547		
Labor market						
Log hourly wage	3.69	0.64	3.69	3.69	0.002	
Weekly hours worked	53.59	13.09	53.23	54.28	-1.05***	
Observations	31,535		20,896	10,639		
Primary activity status						
Regular work	0.18	0.39	0.17	0.22	-0.05***	
Casual labor	0.10	0.30	0.09	0.11	-0.02***	
Self-employment	0.13	0.34	0.13	0.13	0.003*	
Unpaid family work	0.08	0.27	0.08	0.09	-0.009**	
Domestic duties	0.35	0.48	0.35	0.34	0.01***	
Observations	77,737		54,190	23,547		

Notes : The following categories *not literate*, *literate without formal schooling* and *literate below primary* are grouped into the category *below primary.* ♠ Among those individuals who did not complete primary schooling, approximately 76% did not attend school at all and 24% are primary school dropouts. ♣ This category includes all the categories that fall into higher education : *Diploma/certificate course*, *graduate*, *postgraduate and above*. Columns (1) and (2) show the mean and standard deviation of the main variables for the entire sample. Columns (3)-(5) distinguish between individuals with zero childhood exposure to storms from those with positive exposure. Column (5) tests whether means statistically differ from each other across these two groups of individuals. The variable *gender* is equal to one for female individuals. *First born* is equal to one for first born individuals. Finally, *Hinduism* is equal to one for Hindus. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

6 BASELINE RESULTS

6.1 Educational delay

Specification

We evaluate the impact of early childhood exposure to storms on individuals' educational delay using the following specification :

$$Y_i = \alpha_0 + \alpha_1 C_{bd} + \mathbf{X_i'}\boldsymbol{\beta} + \delta_d + \delta_b + \epsilon_i, \tag{4.2}$$

where i is an individual's subscript. Y_i captures an individual's educational delay, measured either using the number of years of educational delay or a dummy variable taking the value of one in the case of an educational delay of at least one year. $\mathbf{X_i'}$ is a vector of individual characteristics including dummy variables indicating if the individual is a female, a first-born child and Hindu respectively. While we drop subscripts where possible, it is understood that $i = (b, d)$, where b denotes the birth-year cohort and d the district. [66]

δ_d is a set of district FE to control for fixed district characteristics that may affect the education level of individuals. δ_b is a set of birth-year cohort FE. The inclusion of both district and birth-year cohort FE implies that identification is achieved using two sources of data variation, i.e. by comparing the educational delays of first, cohorts with different levels of exposure within districts, and second, the same birth-year cohort across districts facing differential exposures. Finally, ϵ_i is the error term.

66. We specifically do not include in the regression controls such as being the household head, married, living in a rural area and for the size of the individual's household. In principle, each of these variables could be affected by childhood exposure to storm and, in consequence, be a bad control.

Importantly, note that our sample only contains individuals who were actually enrolled in the schooling system. Therefore, the effect on the most vulnerable children (e.g. those belonging to scheduled castes who most likely did not attend school at all) is not captured by α_1.

Results

Table 4.3 shows results for equation (4.2). Standard errors are clustered at the state level. Clustering at the state level accounts for the fact that the largest part of funding for education and the coordination of education programs are administered at the state level.[67] In addition, state-level clustering also takes into consideration spatial correlations within state as well as time correlations in the exposure index that result from the fact that the same storm affects multiple birth-year cohorts simultaneously.

In Panel A we measure educational delay as the difference between reported years of schooling and the minimum number of years corresponding to the reported educational attainment. Column (1) shows the baseline result and suggests that exposure to storms leads to a statistically significant delay in completing a given level of education. The estimate indicates that a child with unit exposure will be delayed by 0.31 years on average, which amounts to a delay of approximately 13 weeks (3 months). The left panel of Figure 4.2 indicates that while unit values in the exposure index are exceptional in our sample period, they are actually observed in Orissa. These values result from the 1999 BOB 06 super cyclone (rather than a series of small-scale storms during compulsory schooling), the most intense and destructive tropical cyclone recorded over the period 1990-2000 in India. Extremely severe cyclonic storms are certainly rarer, yet they have been observed again recently, e.g. storm Phailin and Fani made landfall in Orissa in 2013 and 2019 respectively, and super cyclone Amphan hit West Bengal in 2020. In 2021, the extremely severe cyclonic storm Tauktae caused severe damages in Gujarat and ten days later, the severe cyclonic storm Yaas led to grave destruction in West Bengal and Orissa. In light of the recent surge of extremely severe cyclonic events, it is important to provide an interpretation of our estimates for large (unit) values

67. `http://countrystudies.us/india/37.htm`

of the exposure index, as it informs on the educational long-term delays that the current generation of school-attending kids may suffer. If instead we use average exposure in the sample to interpret our results, we obtain an educational delay of roughly six and a half school days. [68]

In columns (2)-(4) we use different sets of FE to investigate the robustness of our results. In column (2) we include state trends to control for trends in state-level education policies and regional disparities in economic growth. Results are quantitatively similar to the baseline estimate. In column (3) we incorporate state-(birth)-year FE, thereby allowing state-level economic conditions at the time of birth of a cohort to affect long-term educational delays. Although the estimate is imprecise, it is worth noting that, even in this demanding specification, the size of the estimate remains nearly identical to the baseline.

In the last column we add state-period FE. The period FE used in the interaction term include three periods : 1985, 1986-1991 and post 1991. This additional set of FE accounts, at least to some extent, for the introduction of the new National Policy on Education introduced under the government of Rajiv Gandhi in 1986, and for its amendment in 1992. An important constituent of the policy is that it called for fulfilling compulsory education for all children up to the age of 14. Although it was introduced at the national level, the policy got effectively adopted at the state level. Results remain similar to the baseline estimates.

In Panel B, we run a linear probability model to explore the effect of childhood exposure to storms on the probability of having an educational delay of at least on year. The estimates obtained are considerably stable across specifications and suggest that being exposed to storms during compulsory schooling increases the probability of a delay. Focusing on the baseline specification (column 1) and examining the impact of a unit exposure (which is likely to be driven by severe events, as describe earlier), the estimate implies an increase of 18 percentage points of the probability of accumulating a delay (i.e. repeating a year or dropping out). Meanwhile, the average

68. Given an average storm exposure of 0.1, 42 weeks per year and assuming that the benchmark school week contains 5 days, this number is computed as $0.31 \cdot 0.1 \cdot 42 \cdot 5$.

exposure increases this probability by 1.8 percentage points. To give a sense of the importance of these findings, consider the states that never experienced storms. Among these states, the share of individuals with an educational delay of at least one year is 0.26. From the baseline estimate in Panel B we can infer that this share would increase by 69% (to a share of 0.44) in the case of an extreme cyclonic storm exposure and by 6.9% (to a share of 0.278) in the case of an average exposure. Hence, these are sizable changes in the probability of being delayed.

Table 4.3:
Educational delay

	Educational delay			
	(1)	(2)	(3)	(4)
Panel A :				
# of years				
Childhood exposure	0.31***	0.27**	0.20	0.29***
	(0.079)	(0.13)	(0.15)	(0.097)
Panel B :				
yes=1, no=0				
Childhood exposure	0.18***	0.20***	0.20**	0.19***
	(0.052)	(0.067)	(0.082)	(0.065)
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes
State trends	No	Yes	No	No
State-(birth)-year FE	No	No	Yes	No
State-period FE	No	No	No	Yes
Observations	77,737	77,737	77,737	77,737

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. In Panel A educational delay is computed as the difference between reported years of schooling and the minimum number of years corresponding to the reported educational attainment. In Panel B, the measure of educational delay is given by a dummy variable taking the value of one in the case of an educational delay of at least one year. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively. Period FE used in the interaction terms include three FE corresponding to the year 1985, the periods 1986-1991 and post 1991.

6.2 Educational attainment

In Table 4.4 we examine whether, in addition to creating educational delays, storms also impact the probability of completing a given level of education. To answer this question we run an ordered logit estimation using a categorical variable indicating the reported educational attainment (0=no formal schooling, 1=primary school, 2=middle school, 3=secondary education, 4=above-secondary education). As in equation (4.2), we include individual controls, a set of district and birth-year FE and cluster standard errors at the state level.

The first column of the table shows the ordered logit estimates. In column (2)-(6) we report the marginal effects of childhood exposure to storms for each category of schooling. The estimates, which are all statistically significant, yield percentage points changes in the probability of completing at most a given category of schooling in the case of unit childhood exposure to storms. As a general statement, positive exposure increases the probability of having no formal schooling and of completing at most primary or middle school. Concurrently, exposure to storms reduces the probability of attaining secondary or post-secondary education.

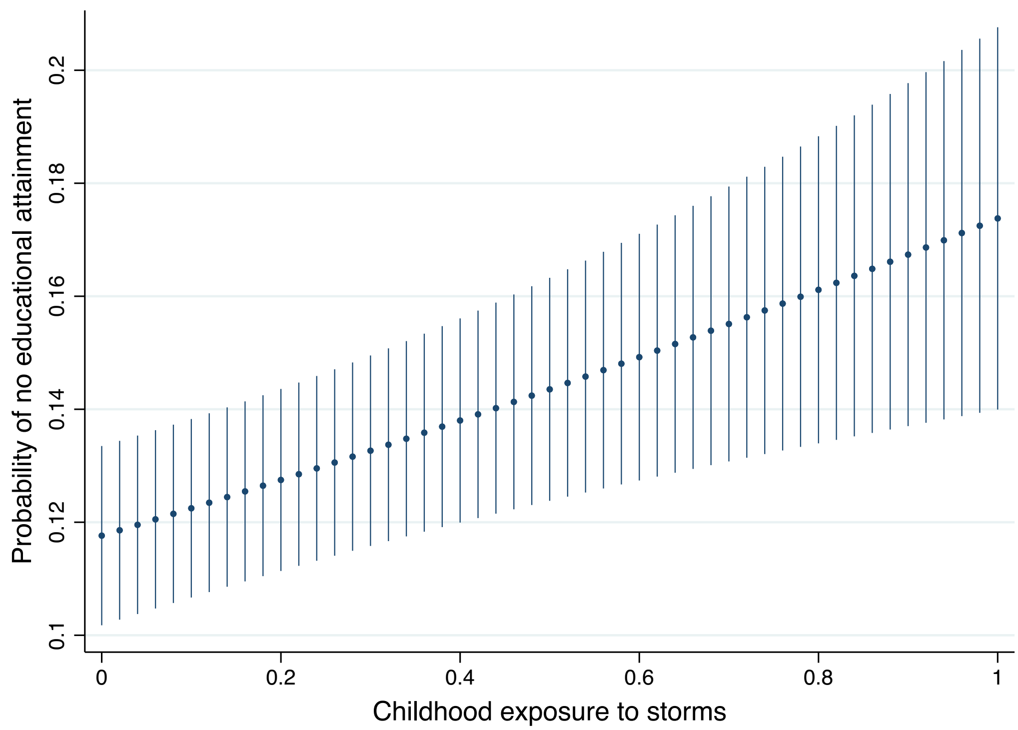
The estimates are the largest for the highest and lowest schooling categories, implying a fall of 8.1 and an increase of 4.8 percentage points in the probability of completing post-secondary education and no formal schooling in the case of unit exposure respectively. Consider for instance kids who were exposed to the 1999 BOB 06 super cyclone during years of compulsory schooling (i.e. $C_{bd} = 1$). Our sample indicates that in Orissa about 17% of individuals have gone past secondary education. Taking this number as benchmark, the estimate translates into a 47% decrease in the probability of completing post-secondary education. If instead we consider the average exposure of 0.1, the decrease would be much smaller and reduce to 4.75%. In the same state, the share of individuals with no formal schooling is 13%. Hence, one can infer that exposure to the 1999 super cyclone led to a 37% increase of individuals lacking basic education, and, as a consequence, a substantial drop in literacy and numeracy rates. Average exposures instead led to a 3.7% increase, which is smaller but still sizable.

In Figures 4.3 to 4.7 we use the estimates from the ordered logit and plot the predicted probabilities of attaining a given category of schooling over the interval [0, 1] of exposures to storms, and their 95% confidence bands, using as benchmark the share of individuals (in the full sample) that belong to a given category. For instance in Figure 4.3, the probability corresponding to an exposure of zero is the share of individuals with no formal education (about 12%, which is in line with the summary statistics in Table 4.2). Overall, the results from this exercise clearly indicate that storms shift the distribution of educational attainment to the left, which is particularly worrisome for developing countries like India whose distribution of skills is already quite skewed to the left.

Table 4.4:
Educational attainment

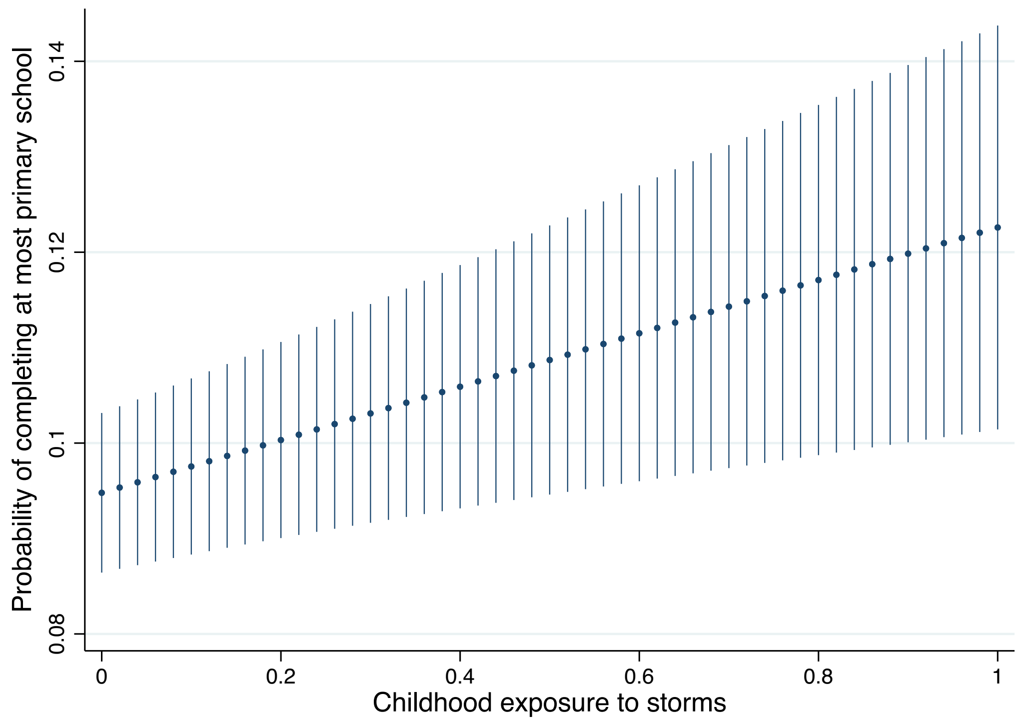
	Logit estimates	No formal schooling	Primary school	Middle school	Secondary education	Above-secondary education
	(1)	(2)	(3)	(4)	(5)	(6)
Childhood exposure	-0.48***	0.048***	0.028***	0.031***	-0.026***	-0.081***
	(0.130)	(0.012)	(0.008)	(0.009)	(0.007)	(0.022)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	77,737	77,737	77,737	77,737	77,737	77,737

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Column (1) shows the results from an ordered logit estimation where the dependent variable is a categorical variable indicating the reported educational attainment (0=no formal schooling, 1=primary school, 2=middle school, 3=secondary education, 4=above-secondary education). Columns (2)-(6) report the marginal effects of childhood exposure to storms for each category of schooling. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively.



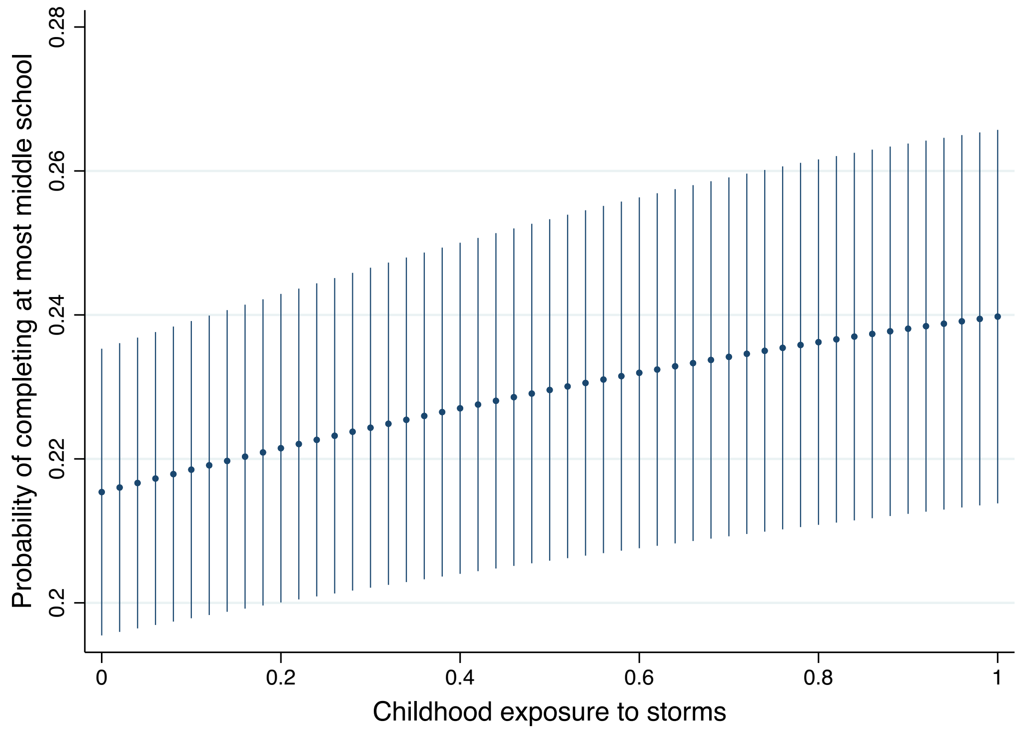
Note : The figure shows the predicted probabilities (and their 95% confidence intervals) of no educational attainment, over the interval [0,1] of exposures, using as benchmark the share of individuals that belong to a given category. Estimates are obtained from Table 4.4, an ordered logit estimation of educational attainment.

FIGURE 4.3 Effect of childhood exposure on the probability of no educational attainment



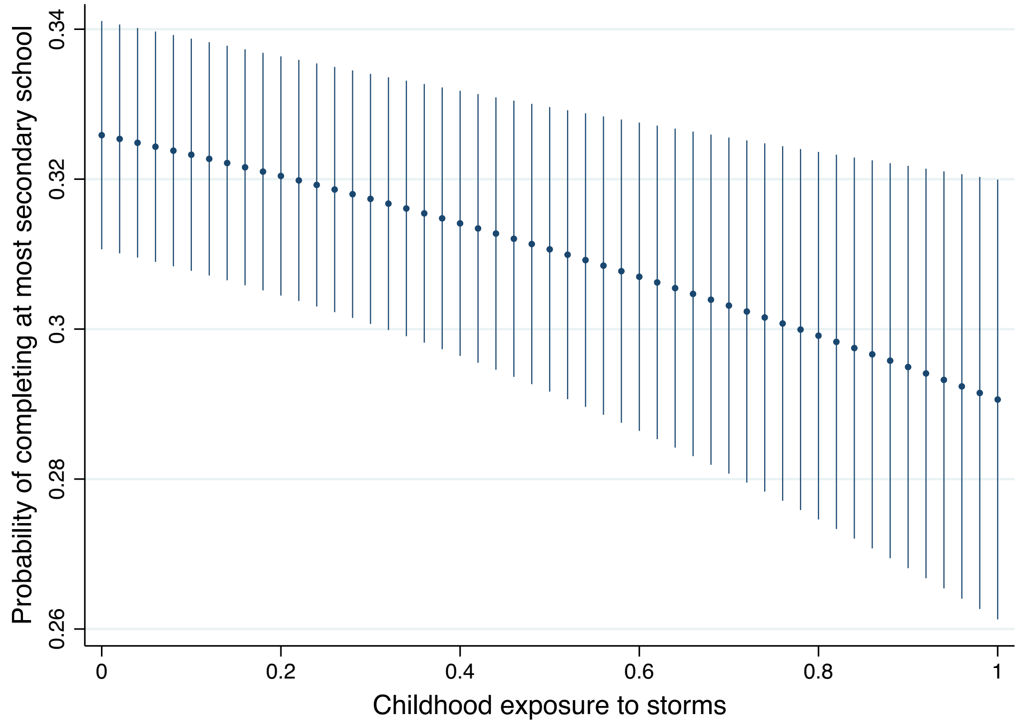
Note : The figure shows the predicted probabilities (and their 95% confidence intervals) of attaining at most primary school, over the interval [0,1] of exposures, using as benchmark the share of individuals that belong to a given category. Estimates are obtained from Table 4.4, an ordered logit estimation of educational attainment.

FIGURE 4.4 Effect of childhood exposure on the probability of attaining at most primary school



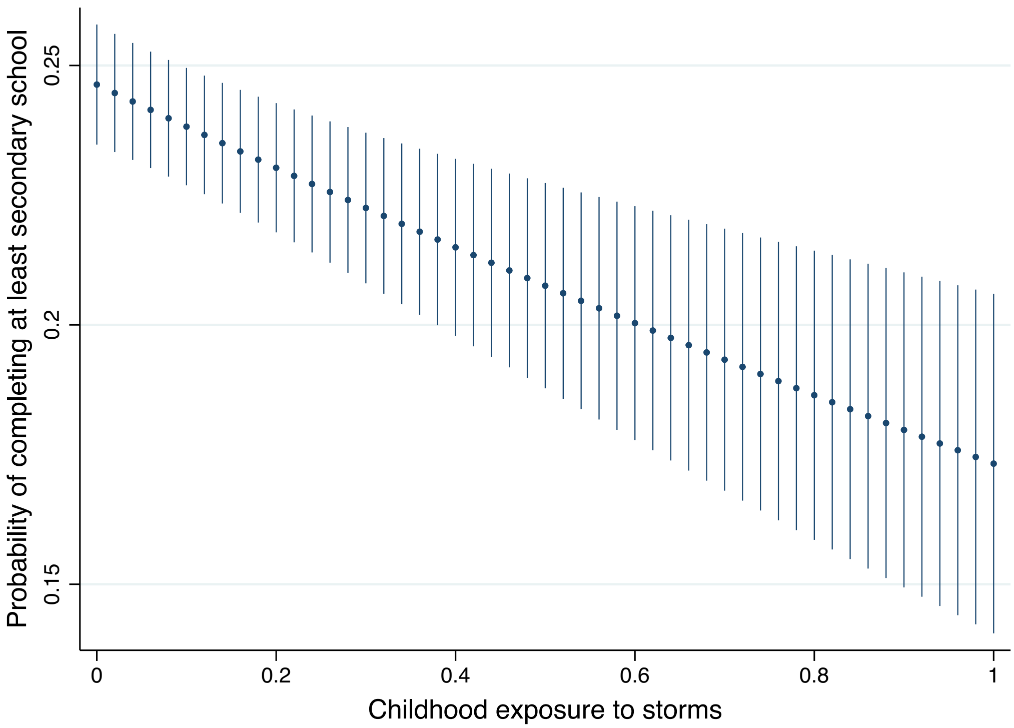
Note : The figure shows the predicted probabilities (and their 95% confidence intervals) of attaining at most middle school, over the interval [0,1] of exposures, using as benchmark the share of individuals that belong to a given category. Estimates are obtained from Table 4.4, an ordered logit estimation of educational attainment.

Figure 4.5 Effect of childhood exposure on the probability of attaining at most middle school



Note : The figure shows the predicted probabilities (and their 95% confidence intervals) of attaining at most secondary school, over the interval [0,1] of exposures, using as benchmark the share of individuals that belong to a given category. Estimates are obtained from Table 4.4, an ordered logit estimation of educational attainment.

Figure 4.6 Effect of childhood exposure on the probability of attaining at most secondary school



Note : The figure shows the predicted probabilities (and their 95% confidence intervals) of attaining at least secondary school, over the interval [0,1] of exposures, using as benchmark the share of individuals that belong to a given category. Estimates are obtained from Table 4.4, an ordered logit estimation of educational attainment.

FIGURE 4.7 Effect of childhood exposure on the probability of attaining at least secondary school

6.3 Type of activity

We expect this educational disruption to be reflected in the type of labor market activity performed in young adulthood, as certain categories of jobs require higher levels of education or at the very least adequate reading, writing and computing skills. We investigate this issue in Table 4.5 by estimating a reduced-form specification of childhood exposure to storms on an indicator variable for each activity. For instance, in column (1) the dependent variable is a dummy variable equal to 1 if the main activity of individual i is to perform regular work. For each type of activity, we include individual controls as in equation (4.2), district FE, birth-year FE, as well as an additional set of state trends to account for economic tendencies that may impact labor markets across states differentially. Estimates suggest that individuals who were exposed to storms during compulsory schooling are less likely to work as regular salaried worker, to be self-employed and more likely to perform domestic duties. However, we find no statistically significant effect on the likelihood to be a casual worker or an unpaid family worker.

Our results suggest that the long-term effect of childhood exposure to storms on the type of labor market activity is sizable. To give a sense of the magnitudes, consider the long-term labor market impacts associated with the average (positive) exposure in our sample (i.e. C_{bd}=0.1). The estimate in column (1) implies a reduction of the probability of being a regular worker of 1.6 percentage points. Among states which never experienced storms over the period 1990-2010, the average share of individuals performing regular work is 0.16. Taking this number as benchmark, the estimate in column (1) implies a 10% reduction in the probability of being employed as a regular worker. For self-employment, column (3) indicates a probability drop of 1.2 percentage points. Using the average share of self-employed workers among states with zero exposure (i.e. a share of 0.14) as a starting point, this percentage points drop amounts to an 8.5% long-term decline in the probability of being self-employed. Finally, the estimate in column (5) implies a 1.8 percentage points raise in the probability of performing domestic duties as primary activity in young adulthood. This raise translates in a 5% change when taking the share of individuals involved in domestic duties in states with zero exposure (i.e. a share of 0.37) as baseline. These changes are larger for kids who were exposed to stronger storms. For instance, with unit exposures and taking the previous baseline shares, the estimates imply changes of approximately 100%, 86% and 49% for regular work, self-employment and domestic duties respectively.

Table 4.5:
Type of activity

	Regular work	Casual labor	Self-employed	Unpaid family work	Domestic duties
	(1)	(2)	(3)	(4)	(5)
Childhood exposure	-0.16**	-0.059	-0.12***	0.046	0.18**
	(0.069)	(0.049)	(0.020)	(0.045)	(0.073)
Individual controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes	Yes
State trends	Yes	Yes	Yes	Yes	Yes
Observations	77,737	77,737	77,737	77,737	77,737

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively.

Finally, Table 4.6 examines whether positive exposure is associated with lower wages and longer hours of employment. In column (1) the sample is restricted to workers who receive a salary, which explains the drop in sample size. In the second column the sample contains all the individuals who report positive hours of work, including people performing unpaid tasks. We find no evidence that, conditional on being employed as regular labor, childhood exposure to storms has permanent effects on wages. By disrupting the education of children, storms are likely to exacerbate income and social inequalities across districts and cohorts in the long run. This increase in inequalities occurs not so much through a reduction in wages but through a change in qualifications and type of activity, inducing less regular jobs and relatively more unpaid work (e.g. domestic duties) which do not provide the social security that formal jobs can supply. Disparities along the income distribution will also widen as individuals with the largest delays tend to belong to particularly vulnerable social groups.

Table 4.6:
Wages and hours of work

	Log hourly wages	Hours of work
	(1)	(2)
Childhood exposure	-0.021	5.72
	(0.19)	(4.55)
Individual controls	Yes	Yes
District FE	Yes	Yes
Birth-year FE	Yes	Yes
State trends	Yes	Yes
Observations	31,534	31,534

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively.

7 ROBUSTNESS

In this section we propose a series of robustness checks involving two falsification tests, the addition of educational controls, the removal of extreme exposures and finally the use of alternative measures of storm exposure. In each case, we start with the discussion of the results for education and then briefly present the estimates obtained on the type of activities performed by individuals.

7.1 Falsification tests

Table 4.7 presents the results of two falsification tests on the educational delay and attainment regressions. In Panel A educational delay is measured by the number of years of delay, while in Panel B it is measured by a dummy variable taking the value of one in the case of an educational delay of at least one year. In Panel C we report the results of ordered logit estimations on educational attainment.

The first falsification test consists in a placebo exercise in which we replace the index of childhood exposure to storms in equation (4.2) with a random exposure obtained by reshuffling C_{bd} over the entire sample. We perform this operation 1000 times and in columns (1)-(3) report the share of replications that produce statistically significant estimates at the 1%, 5% and 10% levels respectively. Overall, results suggest that our coefficients do not capture spurious correlations. The numbers in column (1) indicate that in only 1.2% (Panel C) to 2.6% (Panel A) of the cases, the randomization produces estimates which are statistically significant at the 1% level. Not surprisingly, this share increases when considering higher levels of statistical significance, reaching maximums of 7.9% (Panels A and B) at the 5% level and 13.2% (Panel B) at the 10% level.

In column (4) of the same table we perform a second falsification test. We assign our exposure index to cohorts 10 years older (which were not included in our sample initially). One would expect educational attainments of older cohorts to be unaffected by the occurrence of future storms. Specifically, for each birth-year cohort and district, we assign the value of C_{bd} to the cohort born 10 years earlier. We then examine the effect of this artificial exposure measure on the educational delay of the cohorts born between 1975 and 1985. Regardless of the measure of education used, the estimates obtained are highly statistically insignificant.

In Table 4.8 we repeat these two falsification tests on the type of activities performed by individuals. The results are in line with the ones obtained in Table 4.7, leading us to the same conclusions that our baseline estimates do not capture spurious relationships.

Table 4.7:
Falsification tests, education

	Placebo: Share of estimations with statistical significance at : 1%	5%	10 %	Older cohort assignment + 10 years
	(1)	(2)	(3)	(4)
Panel A :				
Educ. delay : # of years				
Childhood exposure	0.026	0.079	0.128	-0.040
				(0.12)
Panel B :				
Educ. delay : yes=1, no=0				
Childhood exposure	0.024	0.079	0.132	-0.038
				(0.064)
Panel C :				
Educ. attainment				
Childhood exposure	0.012	0.060	0.108	-0.370
				(0.410)
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes
Observations	77,737	77,737	77,737	67,765

Notes : In Panel A educational delay is computed as the difference between reported years of schooling and the minimum number of years corresponding to the reported educational attainment. In Panel B the measure of educational delay is given by a dummy variable taking the value of one in the case of an educational delay of at least one year. In Panel C educational attainment is a categorical variable indicating the reported educational attainment (0=no formal schooling, 1=primary school, 2=middle school, 3=secondary education, 4=above-secondary education). Columns (1)-(3) show the share of statistically significant results over 1000 randomizations, where the childhood exposure measure is randomized over the entire sample. Column (4) shows the estimates obtained using the synthetic index of childhood exposure over the sample consisting of cohorts born between 1975 and 1985. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively.

Table 4.8:
Falsification tests, type of activities

	Placebo Share of estimations with statistical significance at : 1%	5%	10 %	**Older cohort assignment** + 10 years
	(1)	(2)	(3)	(4)
Panel A :				
Regular work				
Childhood exposure	0.016	0.061	0.131	-0.053
				(0.058)
Panel B :				
Casual work				
Childhood exposure	0.016	0.085	0.142	-0.058
				(0.047)
Panel C :				
Self-employed				
Childhood exposure	0.024	0.089	0.144	0.097
				(0.080)
Panel D :				
Unpaid family work				
Childhood exposure	0.023	0.074	0.131	0.022
				(0.032)
Panel E :				
Domestic duties				
Childhood exposure	0.028	0.093	0.146	0.032
				(0.054)
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes
State trends	Yes	Yes	Yes	Yes
Observations	77,737	77,737	77,737	67,765

Notes : Notes : In Panels A-E, the dependent variable is a dummy taking the value of one if an individual's primary activity is regular work, casual work, self-employment, unpaid family work or domestic duties respectively. Columns (1)-(3) show the share of statistically significant results over 1000 randomizations, where the childhood exposure measure is randomized over the entire sample. Column (4) shows the estimates obtained using the synthetic index of childhood exposure over the sample consisting of cohorts born between 1975 and 1985. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively.

7.2 Education controls

The summary statistics in Table 4.2 indicate that, on average, educational delays differ across categories of schooling. In particular, the table shows that while delays are pretty high among individuals who completed at most primary schooling, they tend to decrease as we move to higher educational categories. Hence, it appears that within educational categories, individuals may share observable characteristics or have similar abilities which make them more or less likely to accumulate a delay. Accounting for this fact is challenging given that exposure to storms impacts both educational delay and educational attainment. While using educational categories fixed effects would cause a bad-control problem, we propose two alternative ways of approaching this issue, acknowledging that each approach has its own limitations.

First, we proxy educational attainment using the predicted probability of completing the reported level of education, conditional on observable individual characteristics. Specifically, for each education category we run a linear probability model on a set of individual characteristics : the individual's gender, year of birth and whether she is a first born child and Hindu. Interactions of these variables are also included in each regression. To avoid a bad control issue in the final regression, we focus on the sub-sample of states with zero exposure over the period 1990-2010 and restrict ourselves to individual characteristics that are unlikely to be affected by storms, which admittedly will generate a proxy based on a limited set of observable characteristics. We then use the estimates to predict the probability of completing each level of education for each individual in the full (baseline) sample and use the probability corresponding to the reported level of education as a proxy for educational attainment in our final regression. The result from this exercise is presented in column (2) of Table 4.9 and is exceptionally close to the baseline results (column 1, for comparability purposes).

Second, we include fixed effects capturing parental education. The latter has been shown to be an important predictor of children educational achievements [Björklund and Salvanes, 2011, Guryan, Hurst, and Kearney, 2008, Kim, Tong, and Sun, 2021] but is far less likely to be affected by children exposure to storms. While it is rather implausible that parental education overlaps with children compulsory schooling at low levels of education, one should keep in mind that it may nevertheless be possible that parents enrolled at university have kids attending primary school. This may be particularly true for relatively young parents. One disadvantage of this approach is that parental education is only observed in our sample if both the individual and the parents still live in the same household. Moreover, as married women tend to move in with their husbands' family, the sample will reduce to a sub-sample that contains relatively more males than in the baseline. For this reason, we start by replicating the baseline approach on the sub-sample for which parental education is available (column 3 of the table). Although the sample only represents 42% of the initial sample, the coefficients are very similar to the baseline estimates. In column (4) we proceed to include parental FE and obtain nearly identical estimates.

In Table 4.10 we also perform these two exercises on the type of activities. Panel A shows the baseline results and Panel B includes the predicted probability of completing the reported level of education. Including this predicted probability does not alter the baseline estimates ; exposure to storms during compulsory schooling reduces the probability of being employed as a regular worker, be self-employed and increases the likelihood of performing domestic duties in a statistically significant manner. Panel C replicates the baseline regression on the sub-sample that includes information on parental education. When estimated over the sub-sample the probability to be self-employed appears to be no longer affected by childhood exposure to storms. Importantly, however, including parental education (Panel D) as control produces estimates which are very similar to those obtained in Panel C.

Table 4.9:
Educational controls, education

	Baseline	Predicted educ. attainment	Sub-sample	Parental education
	(1)	(2)	(3)	(4)
Panel A :				
Educ. delay : # of years				
Childhood exposure	0.31***	0.31***	0.36**	0.36**
	(0.079)	(0.079)	(0.15)	(0.15)
Panel B :				
Educ. delay : yes=1, no=0				
Childhood exposure	0.18***	0.18***	0.21***	0.21***
	(0.052)	(0.051)	(0.075)	(0.075)
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes
Predicted educ. attainment	No	Yes	No	No
Parental education	No	No	No	Yes
Observations	77,737	77,737	32,581	32,581

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively. In Panel A educational delay is computed as the difference between reported years of schooling and the minimum number of years corresponding to the reported educational attainment. In Panel B the measure of educational delay is given by a dummy variable taking the value of one in the case of an educational delay of at least one year.

TABLE 4.10:
Educational controls, type of activities

	Regular work	Casual labor	Self-employed	Unpaid family work	Domestic duties
	(1)	(2)	(3)	(4)	(5)
Panel A :					
Baseline					
Childhood exposure	-0.16**	-0.059	-0.12***	0.046	0.18**
	(0.069)	(0.049)	(0.020)	(0.045)	(0.073)
Observations	77,737	77,737	77,737	77,737	77,737
Panel B :					
Predicted educ. attainment					
Childhood exposure	-0.15**	-0.055	-0.12***	0.044	0.18**
	(0.069)	(0.050)	(0.020)	(0.045)	(0.073)
Predicted educ. attainment	Yes	Yes	Yes	Yes	Yes
Observations	77,737	77,737	77,737	77,737	77,737
Panel C :					
Sub-sample					
Childhood exposure	-0.21**	-0.0094	-0.11	0.067	0.12***
	(0.10)	(0.089)	(0.070)	(0.087)	(0.040)
Observations	32,581	32,581	32,581	32,581	32,581
Panel D :					
Parental education					
Childhood exposure	-0.21*	0.00017	-0.11	0.069	0.13***
	(0.11)	(0.088)	(0.074)	(0.087)	(0.039)
Parental education	Yes	Yes	Yes	Yes	Yes
Observations	32,581	32,581	32,581	32,581	32,581
Individual controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes	Yes
State trends	Yes	Yes	Yes	Yes	Yes

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively. Panel A shows baseline estimates. Panel B includes the predicted probability of completing the reported level of education. Panel C replicates the baseline specification on the sub-sample for which parental education is available. Panel D controls for parental education.

7.3 Excluding extreme values

In Table 4.11 we evaluate the sensitivity of our results to extreme values of exposure. The panel structure of the table is similar to Table 4.7. Column (1) shows the baseline results. In column (2) we exclude individuals located in Orissa, which, as discussed earlier, exhibits exceptionally large values of exposures due to the 1999 super cyclone BOB 06. Results obtained on this sub-sample are not different from the baseline estimates. In column (3) we interact the exposure index and a dummy taking the value of one for individuals who live in Orissa and obtain estimates which align with those obtained on the sub-sample. Hence, it appears that our baseline results are not driven by the super cyclonic event that took place in Orissa in 1999. Finally, in the last column we consider all severe events and recompute the exposure index, removing all the winds with values falling above the 95th percentile of the wind distribution. Not surprisingly, taking out extremes causes the estimates to shrink, albeit by a relatively small extent. The estimates remain qualitatively similar and precise.

In Table 4.12 we perform a similar exercise for the type of activities performed by individuals as they reach young adulthood. The column structure of the table is identical to Table 4.5, Panel A reports the baseline results, Panel B and C examine whether the baseline results are driven by observations in the state of Orissa and the last panel removes severe winds. The conclusions we achieve are in line with those obtained in Table 4.11, hence our results are not sensitive to extreme values of exposure.

Table 4.11:
Sensitivity to extreme values of exposure, education

	Baseline	Excluding Orissa	Interaction : Orissa × C_{bd}	Excluding extreme winds
	(1)	(2)	(3)	(4)
<u>Panel A :</u>				
Educ. delay : # of years				
Childhood exposure (C_{bd})	0.31***	0.30***	0.30***	0.20***
	(0.079)	(0.10)	(0.10)	(0.067)
Orissa × C_{bd}			0.016	
			(0.10)	
<u>Panel B :</u>				
Educ. delay : yes=1, no=0				
Childhood exposure (C_{bd})	0.18***	0.16**	0.16**	0.10**
	(0.052)	(0.061)	(0.061)	(0.040)
Orissa × C_{bd}			0.055	
			(0.061)	
<u>Panel C :</u>				
Educ. attainment				
Childhood exposure (C_{bd})	-0.48***	-0.53***	-0.55***	-0.32***
	(0.13)	(0.18)	(0.19)	(0.10)
Orissa × C_{bd}			0.23	
			(0.19)	
Individual controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes
Observations	77,737	75,192	77,737	77,737

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively. In Panel A educational delay is computed as the difference between reported years of schooling and the minimum number of years corresponding to the reported educational attainment. In Panel B the measure of educational delay is given by a dummy variable taking the value of one in the case of an educational delay of at least one year. In Panel C educational attainment is a categorical variable indicating the reported educational attainment (0=no formal schooling, 1=primary school, 2=middle school, 3=secondary education, 4=above-secondary education). C_{bd} in the interaction term denotes childhood exposure to storms.

TABLE 4.12:
Sensitivity to extreme values of exposure, type of activities

	Regular work	Casual labor	Self-employed	Unpaid family work	Domestic duties
	(1)	(2)	(3)	(4)	(5)
Panel A :					
Baseline					
Childhood exposure (C_{bd})	-0.16**	-0.059	-0.12***	0.046	0.18**
	(0.069)	(0.049)	(0.020)	(0.045)	(0.073)
Observations	77,737	77,737	77,737	77,737	77,737
Panel B :					
Excluding Orissa					
Childhood exposure (C_{bd})	-0.23***	-0.088	-0.099***	0.069	0.18*
	(0.039)	(0.062)	(0.020)	(0.065)	(0.10)
Observations	75,192	75,192	75,192	75,192	75,192
Panel C :					
Interaction : Orissa $\times$ C_{bd}					
Childhood exposure (C_{bd})	-0.23***	-0.087	-0.10***	0.070	0.18*
	(0.039)	(0.063)	(0.019)	(0.065)	(0.10)
Orissa $\times$ C_{bd}	0.23***	0.092	-0.061***	-0.080	-0.025
	(0.041)	(0.067)	(0.018)	(0.066)	(0.099)
Observations	77,737	77,737	77,737	77,737	77,737
Panel D :					
Excluding extremes winds					
Childhood exposure C_{bd}	-0.13***	-0.059	-0.072***	0.046	0.12**
	(0.041)	(0.041)	(0.021)	(0.035)	(0.056)
Observations	77,737	77,737	77,737	77,737	77,737
Individual controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes	Yes
State trends	Yes	Yes	Yes	Yes	Yes

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively. C_{bd} in the interaction term denotes childhood exposure to storms.

7.4 Alternative measures of childhood exposure to storms

In Tables 4.13 and 4.14 we experiment with alternative specifications of C_{bd}. First, we change the functional form that captures the force exerted by winds on structures (we replace the square with a cube in equation 4.3 of Appendix B). A few studies in the U.S. claim that the energy released by a storm and the force exerted by winds on physical structures are related in a cubic and not a quadratic manner [see the technical HAZUS manual of the Federal Emergency Management

Agency – FEMA – of the U.S. Department of Homeland Security and Emanuel, 2005]. We account for this fact in column (2) and (4) of Table 4.13 and in Panels B and D of Table 4.14, where we use a cubic specification.

Second, we alter the threshold above which a wind is classified as a storm. Throughout the paper, we define the benchmark windspeed threshold likely to generate damages at 50 knots, following Emanuel [2011]. In columns (3) and (4) of Table 4.13 and in Panels C and D of Table 4.14, we move the storm threshold up to 64 knots (corresponding to a category 1 cyclone on the Saffir-Simpson scale). In column (5) of table 4.13 and in Panel E of Table 4.14 we completely drop the notion of threshold and use all the winds sweeping the country during a cyclone.

Finally, in column (6) of Table 4.13 and in Panel F of Table 4.14, we use the HURRECON model (see Appendix B.2 for more details) in order to compute the maximum windspeed hitting each district, following Boose, Serrano, and Foster [2004] instead of Deppermann [1947].

TABLE 4.13:
Alternative measures, education

	Baseline	50, cubic	64, square	64, cubic	All winds	HURRECON
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A :						
Educ. delay : # of years						
Childhood exposure	0.31***	0.47***	0.42***	0.50***	0.0033	0.33***
	(0.079)	(0.12)	(0.094)	(0.17)	(0.0079)	(0.090)
Panel B :						
Educ. delay : yes=1, no=0						
Childhood exposure	0.18***	0.28***	0.25***	0.31***	0.0006	0.21***
	(0.052)	(0.052)	(0.049)	(0.070)	(0.0042)	(0.054)
Panel C :						
Educ. attainment						
Childhood exposure	-0.48***	-0.60***	-0.58***	-0.54***	0.006	-0.60***
	(0.13)	(0.20)	(0.15)	(0.21)	(0.013)	(0.15)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	77,737	77,737	77,737	77,737	77,737	77,737

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively. In Panel A educational delay is computed as the difference between reported years of schooling and the minimum number of years corresponding to the reported educational attainment. In Panel B the measure of educational delay is given by a dummy variable taking the value of one in the case of an educational delay of at least one year. In Panel C educational attainment is a categorical variable indicating the reported educational attainment (0=no formal schooling, 1=primary school, 2=middle school, 3=secondary education, 4=above-secondary education).

In each column of Table 4.13 the estimated coefficients remain positive and precisely estimated (with the exception of column 5) and, as expected, the magnitudes of the estimates become larger as the threshold used to compute district exposure increases to capture less frequent but more powerful winds. Using cubic specifications also tend to inflate the estimates. For instance, in column (3) of Panel A, where district exposure is computed using a threshold of 64 and a quadratic function, the estimate grows to 0.42 against 0.31 in the baseline (column 1). This number corresponds to an increase in delay of 8.8 school days for the average exposure index and of 17.7 weeks in districts with the highest exposure. The corresponding estimate in Panel B implies that in these districts, individuals are 25 percentage points more likely to have an educational delay of one year or more, in comparison to storm-sheltered places. The same pattern is observed for the ordered logit estimates in Panel C, the magnitude is slightly bigger but signs and statistical significance are left untouched. Including all winds when computing storm exposure (column 5) introduces a downward bias on our estimates and causes them to be statistically insignificant. This change is not surprising as accounting for all winds means attributing non-zero childhood exposure values to many districts that were only exposed to very mild windspeeds.

Table 4.14:
Alternative measures, type of activities

	Regular work	Casual labor	Self-employed	Unpaid family work	Domestic duties
	(1)	(2)	(3)	(4)	(5)
Panel A :					
Baseline					
Childhood exposure	-0.16**	-0.059	-0.12***	0.046	0.18**
	(0.069)	(0.049)	(0.020)	(0.045)	(0.073)
Panel B :					
50, cubic					
Childhood exposure	-0.16	-0.044	-0.20***	0.036	0.25*
	(0.15)	(0.061)	(0.016)	(0.048)	(0.13)
Panel C :					
64, quadratic					
Childhood exposure	-0.16	-0.045	-0.18***	0.036	0.23**
	(0.12)	(0.056)	(0.0092)	(0.049)	(0.11)
Panel D :					
64, cubic					
Childhood exposure	-0.12	-0.016	-0.23***	0.016	0.25
	(0.16)	(0.049)	(0.042)	(0.030)	(0.16)
Panel E :					
All winds					
Childhood exposure	0.0039	0.0017	-0.011***	-0.0024	0.0012
	(0.0045)	(0.0040)	(0.0037)	(0.0018)	(0.0038)
Panel F :					
HURRECON					
Childhood exposure	-0.16	-0.036	-0.14***	0.031	0.19**
	(0.10)	(0.055)	(0.022)	(0.041)	(0.083)
Observations	77,737	77,737	77,737	77,737	77,737
Individual controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes	Yes
State trends	Yes	Yes	Yes	Yes	Yes

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively. In Panel A we find baseline results. In Panel B the measure of childhood exposure is computed using a threshold of 50 knots and a cube. In Panel C the measure of childhood exposure is computed using a threshold of 64 knots and a square. In Panel D the measure of childhood exposure is computed using a threshold of 64 knots and a cube. In Panel E the measure of childhood exposure is computed using all winds. In Panel F the measure of childhood exposure is computed using the HURRECON model, a threshold of 50 knots and a square.

In Table 4.14 we look at how the impact of childhood exposure on the type of activity changes when we modify the specification of the index. The magnitudes of the results are similar to the baseline but the estimates are less precise. The impact on casual labor, self-employment, unpaid family work and involvement in domestic duties is not affected by the different specifications. Instead, we find that the negative and statistically significant impact on regular work is estimated less precisely when the definition of storm is altered. It is worth mentioning that the coefficient is still borderline significant in some cases : in Panel C the p-value is 0.18 and in Panel F it is 0.11.

8 HETEROGENEITY OF THE EFFECTS

In this section we study how the effects of interest differ across males and females, and urban and rural areas. First, we look at the impact on education across the four groups. Columns (1) and (2) of Table 4.15 show results for educational delay, while columns (3) to (6) present results for educational attainments. The four subsamples have roughly the same size. Results for males (Panel A) are similar in sign and statistical significance to those for females (Panel B). Yet, the table points towards significant differences in terms of the magnitude of the effects. For educational delay, the results for males are roughly 64% larger. These results, even if less pronounced, are similar to what Takasaki [2017] finds regarding boys and girls' school enrolment in the aftermath of a cyclone in the Fiji Islands. This difference in the effect may be due to the different speed of physical development between genders. Often, boys 10 years and older are already physically capable to help their parents in reconstruction activities. If storms were to damage the family farm or fields, a household may decide to keep boys home to help and send girls to school, causing boys to accumulate a delay but not girls. Regarding educational attainment, columns (3) to (6) tell us that storms tend to increase the probability of not having formal education and of completing at most middle school, while they decrease the probability of completing secondary or higher education. While the impact on the educational delay is smaller for females, their probability of not going further than primary school is 178% larger than for males (0.078 vs 0.028) and is statistically significant at the 1% level. These results seem to indicate that while males accumulate an educational delay, females may just end their educational career earlier.

TABLE 4.15:
Education : gender, rural and urban sub-samples

	Educational delay		Category of schooling completed : yes=1, no=0				
	# of years	yes=1, no=0	No educ.	Primary	Middle	Secondary	Above
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A :							
Male							
Childhood exposure	0.41***	0.23***	0.028**	0.025**	0.038**	-0.014**	-0.077**
	(0.13)	(0.061)	(0.011)	(0.010)	(0.015)	(0.0056)	(0.031)
Observations	39,272	39,272	39,272	39,272	39,272	39,272	39,272
Panel B :							
Female							
Childhood exposure	0.25**	0.14***	0.078***	0.031***	0.025***	-0.040***	-0.095***
	(0.12)	(0.049)	(0.028)	(0.012)	(0.0095)	(0.015)	(0.034)
Observations	38,465	38,465	38,465	38,465	38,465	38,465	38465
Panel C :							
Rural							
Childhood exposure	0.27**	0.19***	0.054**	0.038**	0.059**	0.0030**	-0.15**
	(0.10)	(0.060)	(0.024)	(0.016)	(0.025)	(0.0014)	(0.066)
Observations	42,281	42,281	35,456	35,456	35,456	35,456	35456
Panel D :							
Urban							
Childhood exposure	0.33***	0.10**	0.051**	0.025**	0.020**	-0.045**	-0.052**
	(0.12)	(0.046)	(0.022)	(0.011)	(0.0086)	(0.019)	(0.023)
Observations	35,454	35,454	42,281	42,281	42,281	42,281	42,281
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively. In Panel A we present results for the sub-sample of males, and in Panel B, C, and D for females, urban, and rural individuals respectively.

Panel C and D show results for rural and urban areas, respectively. In rural areas, the majority of households are involved in farming. These households are likely to take their children out of school in the aftermath of a storm to perform reconstruction and field activities, generating an educational delay, as found in columns (1) and (2) of Table 4.15. In addition, we find that the probability of completing higher education decreases. Maybe surprisingly, in urban areas, the story seems to be similar. We find an increase in educational delay, and a higher probability of attaining lower educational levels, starting with primary school.

In Table 4.16, we look at what happens to the type of activity undertaken in young adulthood for each of the four subsamples. We find that the probability of being involved in regular work (i.e. high quality salaried jobs) and to be self-employed decreases across the four subgroups, while the probability of performing domestic duties increases across the board.

Table 4.16:
Gender, rural and urban sub-samples, type of activities

	Regular work	Casual labor	Self-employed	Unpaid family work	Domestic duties
	(1)	(2)	(3)	(4)	(5)
Panel A :					
Male					
Childhood exposure	-0.29**	-0.025	0.0036	0.094	0.025*
	(0.14)	(0.082)	(0.056)	(0.068)	(0.013)
Observations	39,272	39,272	39,272	39,272	39,272
Panel B :					
Female					
Childhood exposure	-0.079***	-0.071	-0.20***	0.019	0.29**
	(0.028)	(0.083)	(0.044)	(0.036)	(0.12)
Observations	38,465	38,465	38,465	38,465	38,465
Panel C :					
Rural					
Childhood exposure	-0.098*	-0.062	-0.18***	0.093	0.16**
	(0.054)	(0.058)	(0.019)	(0.065)	(0.074)
Observations	42,281	42,281	42,281	42,281	42,281
Panel D :					
Urban					
Childhood exposure	-0.36***	-0.026	0.028	-0.045*	0.24***
	(0.093)	(0.033)	(0.038)	(0.024)	(0.043)
Observations	35,454	35,454	35,454	35,454	35,454
Individual controls	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Educ.-Birth year FE	Yes	Yes	Yes	Yes	Yes

Notes : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the state level. Individual controls include dummy variables indicating if the individual is a female, the first born child and Hindu respectively. In Panel A we present result for the sub-sample of males, and in Panel B, C, and D for females, urban, and rural individuals respectively.

9 CONCLUSION

In this paper we look at how exposure to storms during compulsory schooling affects long-term educational attainments and the type of activity performed in early adulthood. Using storm data from NOAA, we construct a measure of exposure to storms over the course of compulsory schooling for individuals aged 23 to 33 in the 2018 release of the PLFS. Individuals hit by a storm during these important years tend to accumulate an educational delay and are less likely to complete higher education. We also find that exposure to storms reduces the probability of obtaining a regular salaried job, or being self-employed and increases the probability of performing domestic duties as primary activity in young adulthood.

Duflo [2001] finds that economic returns to education range from 6.8 to 10.6 percent in Indonesia. Applying these numbers to India, our estimates of educational delay imply that a unit exposure to storms during the years of compulsory schooling could cause a lifelong 2.1% to 3.3% fall in returns on average. For the average exposure in our sample, the corresponding lifelong fall in returns ranges between 0.13% and 0.21%.[69] This is an important number considering that storms typically last less than a week. Although we cannot test this formally, our results also suggest that the skill distribution of cohorts subjected to positive exposures during compulsory schooling are more skewed to the left. Our findings consequently hint towards a deskilling of the geographical areas prone to natural catastrophes in the long-term and towards a widening of wage inequality within and across districts. The effects we find are rather moderate for the average exposure over the period 1990-2010 – yet they reach sizable magnitudes for more severe exposures which recently became the norm rather than the exception [Emanuel, 2021].

As we pointed out earlier, there are two main channels through which storms may affect education. On the schooling supply side, a disaster may destroy public infrastructure, like roads and schools, creating punctual delays in schooling due to the impossibility to attend classes

69. A unit exposure to storms causes an average delay of 13 weeks out of the 42 weeks of a school year, which amounts to 31% of the year. Multiplying this number by Duflo [2001]'s estimates we obtain a reduction of 2.1% to 3.3% in returns. For the average exposure, educational delays represent 2% of the year, which translates into a fall of 0.13% to 0.21% in returns.

[see for instance Duflo, 2001, Jaume and Willèn, 2019]. A storm can also impact the demand for schooling, for instance by generating PTSD in children, a condition that has been shown to hinder their scholastic and labor market performance [e.g. Blaikie, Cannon, Davis, and Wisner, 2014, Cutter, Boruff, and Shirley, 2003, Kar and Bastia, 2006, Neria et al., 2008]. Or it may result in a negative income shock for the household, e.g. by destroying crops and farms in rural areas or part of production facilities in both, rural and urban areas [see for instance Pelli et al., 2020]. Negative income shocks can be temporary and last until physical assets are rebuilt, or permanent, if for instance the loss of a season's crop puts a farming household in debt and cripple it financially for years to come.

While it is likely that our findings reflect the interaction of all these different channels, our data do not allow us to identify their relative contribution in increasing educational delays, making it harder to formulate precise policy interventions. Nevertheless, the results from Section 8 suggests that the detrimental effects of storms are present not only in rural but also in urban areas, therefore evoking the need for wide-scale policies aimed at helping both regions. In our opinion however, the set of policies should differ across regions and be adapted to their intrinsic economic activities. In rural areas, we recommend an emergency system of Conditional Cash Transfers (CCT) covering reconstruction and the loss of agricultural and farm income. To limit educational delays, this policy should be conditional on regular and uninterrupted school attendance. In urban areas, individuals are more likely to be laid off as storms destroy firms' facilities and impair production networks. These layoffs and job losses can drag a family down a spiral of poverty especially if unemployment insurance and social assistance programs are dysfunctional, as it is often the case in developing countries. In consequence, when social safety nets are weak, children may be taken out of school earlier, jeopardizing their future employment prospects. A possible policy in urban areas would be to subsidize firms in order for them to keep paying their employees during the reconstruction of their physical infrastructure. These subsidies could be modeled on the policies that have been implemented by many countries during the COVID-19 pandemic and should be complemented with social assistance programs targeted at individuals who are not regular salaried workers.

The results obtained in this paper help us predict part of the long-term impacts of the COVID-19 pandemic that started towards the end of 2019. Government-mandated school closures led to a reduction in the schooling supply, at least temporarily. On the demand side, the literature documents that the pandemic gave rise to PTSD [see for instance Phelps and Sperry, 2020, Yue, Zang, Le, and An, 2020, Zhou, 2020] and, by causing heterogeneous losses of employment, heterogeneous negative income shocks [see Di Pietro, Biagi, Costa, Karpinski, and Mazza, 2020]. In light of our findings, we can infer that the pandemic will likely generate an educational delay for many students, compromise educational attainment and increase the drop-out rate. Consistent with this argument, Kuhfeld, Soland, Tarasawa, Johnson, Ruzek, and Liu [2020] predict that returning students will have only approximately 63-68% of the reading capacity and 37-50% of the mathematical knowledge with respect to a usual school year. Moreover, our results also suggest a change in the type of activity that current students will engage in during early adulthood. This change will likely be reflected in the share of individuals with regular salaried work and also in the type of occupations or tasks individuals will perform. Furthermore, we expect the ensuing fall (increase) in the supply of high-skill (low-skill) workers for the cohorts attending school during 2020-2021 to contribute to increased wage inequality. In light of this evidence, it should be paramount for policymakers to find a way to allow these students to make up for the delay accumulated. In the absence of such an effort, economic inequalities are meant to increase for individuals who were in school during the COVID-19 pandemic.

A APPENDIX. INDIAN SCHOOLING SYSTEM

In India, individuals are subject to compulsory education from grade 1 to grade 8 starting at the age of 6 years old. All students follow a similar path until they complete higher secondary schooling (grade 12). The Indian education system is divided into five stages. Indian children complete lower elementary education in five years, upper elementary education in three years, secondary education in two years and upper secondary in two years. After obtaining their higher secondary education, Indian children have the option to join the labor market or pursue their post secondary education. They have the choice between five post-secondary paths including post-secondary diploma, graduate degree which can be followed by postgraduate education. Post-secondary diploma takes one year, graduate studies about three years and a postgraduate degree an average of two years. The post-secondary diploma is focused on developing technical skills such as nursing or engineering. Often, individuals complete a diploma prior to starting a university degree. Similar to other educational systems, students can study numerous fields at university and need to complete a graduate degree before starting their postgraduate career. In terms of above graduate degree education, a small portion of the Indian population pursue postgraduate education, with the majority completing a masters degree. The first two post-secondary paths consist of obtaining either a diploma/certificate course or a graduate degree. The third postgraduate path is a mix of both where students obtain a technical degree and a graduate education (i.e. : Engineering). The fourth and fifth paths are the continuity of previous path by obtaining a postgraduate degree after completion of their graduate degree or diploma/certificate and graduate degree. Individuals pursuing the postgraduate paths obtain a masters, Ph.D., M.D. or M.B.A. degree. [70]

The Indian government first formulated the National Policy on Education (NPE) in 1968, under the governance of Prime Minister Indira Gandhi. Its main objective was to promote literacy and numeracy in urban and rural areas and to implement compulsory schooling for all children. Importantly, the NPE has a national curricular framework that contains a common core,

70. More details on the educational system of India and how it compares to other systems can be found here : https ://wenr.wes.org/2018/09/education-in-india

such as the learning of Hindi as a common language for all Indians, but allow also flexibilities that take into account region-specific components. For this reason, our empirical strategy adds state-period FE in order to account for regional differences in the implementation of the policy. Given the pace of changing time, the NPE was amended in 1985 by the government of Prime Minister Rajiv Gandhi and adopted by the parliament in august 1986. The new NPE included, among other measures, the spread of adult literacy and ensuring primary education for all [Ray, Satpathy, and Chandra, 2013]. The policy was amended again six years later (august 1992), with the constant objective of improving the quality of the Indian education system. In 2020, the government of Prime Minister Narendra Modi decided to amend the NPE of 1986, with an important change regarding the medium of instruction at school. The new NPE (2020) suggest to use the mother tongue or the local language as the medium of instruction and gives state governments the decision regarding its implementation. [71]

71. https://timesofindia.indiatimes.com/home/education/news/nep-language-policy-broad-guideline-government/articleshow/77272709.cms

B APPENDIX. DISTRICT EXPOSURE AND WIND SPEED

B.1 District exposure to tropical storms

In what follows we describe how we construct x_{dt}, the index of exposure to storms of district d in year t. This index is given by the following quadratic specification :

$$x_{dt} = \sum_{h \in H_t} \frac{(w_{dh} - 50)^2}{(w^{max} - 50)^2} \quad \text{if} \quad w_{dh} > 50, \tag{4.3}$$

where H_t is the set of storms in year t and w_{dh} is the maximum wind speed associated with storm h and to which district d was exposed. We describe the construction of w_{dh} below. The term w^{max} denotes the maximum wind speed observed over the entire sample. In order to capture the force exerted by winds on structures, we assume a quadratic functional form between district exposure to storms and winds, as in Yang [2008] and Pelli and Tschopp [2017]. [72] Given the poor quality of construction materials, infrastructures and buildings in India are already vulnerable at low wind intensities. For these reasons, we focus on a threshold of 50 knots, as in Emanuel [2005], as opposed to 64 knots – the threshold for a category 1 cyclone according to the Saffir-Simpson scale. By definition, $x_{dt} \in (0, \mid H_t \mid)$, with a value of 0 indicating zero district exposure to storms (i.e. winds in district d are below the threshold limit) and with $\mid H_t \mid$ indicating the number of elements (storms) in set H_t.

B.2 Wind speed at the district level

Baseline : the Rankine-combined formula [Deppermann, 1947]

We now turn to the construction of w_{dh}, i.e. the maximum wind speed associated with storm h in district d. The variable is constructed using data from the National Oceanic and Atmospheric Administration (NOAA) Tropical Prediction Center and specifically using storms' best

72. In Section 7.4 we experiment with a variety of alternative specifications of district exposure to storms.

tracks in the North Indian and South Indian basins over the period 1990-2010. Best tracks contains the full history of each storm, with information at 6-hours intervals on the latitude, longitude, date and wind speed at the eye of each storm.

We first linearly interpolate storms' best tracks at every kilometre and obtain, for each interpolated kilometre, a landmark k with a set of coordinates and e_k, the windspeed at the eye of the storm. For each district that falls in the vortex associated with a landmark we use the Rankine-combined formula [Deppermann, 1947] and compute winds at the district's centroid. The formula describes wind fields in the following way :

$$\begin{aligned} w_{dk} &= e_k \cdot \left(\frac{D_{dk}}{26.9978}\right) \text{ if } D_{dk} \leq 26.9978 \\ w_{dk} &= e_k \cdot \left(\frac{26.9978}{D_{dk}}\right)^{0.5} \text{ if } D_{dk} > 26.9978, \end{aligned} \tag{4.4}$$

where D_{dk} is the distance between the centroid of district d and landmark k. The number 26.9978 corresponds to Simpson and Riehl radius of maximum wind speed in knots, i.e. the distance between the eye and the point where wind reaches its maximum speed. [73] Hence, according to this formula, winds first increase exponentially up to a maximum and then, decrease rapidly. Finally, we obtain one measure of windspeed per district and storm by retaining the maximum windspeed to which a district was exposed :

$$w_{dh} = \max_{k \in H_t} \{w_{dk}\}.$$

73. In reality, each cyclone has a different radius of maximum windspeed, which is calculated using the difference in barometric pressure between the center and the outskirts of the storm. Unfortunately, cyclone data are characterized by a high number of missing data when it comes to barometric pressure. For this reason we decided to follow Simpson and Riehl [1981] and Hsu and Zhongde [1998] and apply the average radius of maximum windspeed, 50 km, to all the cyclones considered in this paper.

Alternative specification : the HURRECON model [Boose, Foster, and Fluet, 1994]

The HURRECON model [see Boose et al., 1994, Boose, Chamberlin, and Foster, 2001, Boose et al., 2004] describes sustained wind velocity at any point within a cyclone's vortex using information on the track, size, intensity, and cover type (land or water) of a hurricane. Using this model, we compute sustained wind velocity at each district centroid as follows : [74]

$$w_{dk} = F\left[V_k - S(1-\sin T)\frac{V_f}{2}\right]\left[\left(\frac{R_m}{R}^B e^{1-\left[\frac{R_m}{R}\right]^B}\right)\right]^{1/2} \tag{4.5}$$

where F is a scaling parameter capturing the effect of friction set at 0.8, since all the point of interest to us are situated on land (this parameter is usually set equal to 1 for points over water and to 0.8 for points over land) ; V_k is the wind velocity at the eye at landmark k, which we linearly interpolate from the best track data ; S is a scaling parameter for the asymmetry due to the forward motion of the storm, set to 1 [i.e. peak wind speed on the right side minus peak wind speed on the left side equals the forward velocity of the hurricane – V_f, as defined in Boose et al., 2001] ; T is the clockwise angle between the forward path of the hurricane and a radial line connecting the eye of the hurricane to the population-weighted centroid of a county ; V_f is the forward velocity of the hurricane, i.e. the speed at which the hurricane is moving forward ; R_m is the radius of maximum winds, set as in the previous approach at 26.9978 ; R is the radial (or Euclidean) distance from the center of the hurricane to the population-weighted centroid of a county ; and B is a scaling parameter controlling for the shape of the wind profile curve (usually included between 1.2 and 1.5, and set at 1.35). The parameters of this equation, adapted from Holland's equation for the cyclostrophic wind [Holland, 1980], are set following Boose et al. [2004] that parameterized and validated the model.

74. Velocity and wind direction are measured relative to the surface of the Earth, and angles are measured in degrees.

CHAPITRE 5
CONCLUSION

En résumé, cette thèse a étudié l'effet d'un choc positif de productivité agricole et d'un choc négatif de désastre naturel sur les décisions d'adaptation des ménages en Inde. Les trois chapitres ont utilisé une approche purement empirique basée sur l'évaluation d'impact, avec des estimations par variable instrumentale. L'objectif général visé par cette thèse est principalement de montrer l'effet d'une politique économique ou d'un phénoméne naturel sur les décisions prises par les ménages pour améliorer leur bien-être. Et le choix de l'Inde comme pays d'étude repose sur le fait qu'il représente un modéle à suivre, du fait de son dynamisme économique et de ses caractéristiques démographiques similaires à bon nombre de pays en développement à travers le monde. En tenant compte des imperfections liées à chaque chapitre, cette thèse essaye de montrer en partie la nature de l'interaction entre les individus et l'environnement dans lequel ils évoluent.

Le premier chapitre de cette thèse a étudié l'effet de l'electrification rurale sur le mouvement de la main d'oeuvre entre les secteurs de l'économie, ainsi que son effet sur l'utilisation des ressources naturelles telles que la couverture forestière et la nappe phréatique. Le principal enseignement que ce chapitre a montré est que l'utilisation de l'electricité à des fins d'irrigation augmente la productivité des fateurs de production (travail et terre). Cela permet aux ménages ruraux de réduire les quantités de ces deux facteurs nécessaires à la production agricole. Ainsi, le surplus du facteur travail qui se libére des tâches agricoles est réorienté vers les autres secteurs de l'économie. Ce comportement au niveau des ménages ruraux engendre la transformation structurelle de l'économie Indienne lorsqu'il est agrégé au niveau national. Par la suite, ce chapitre a essayé de voir le "revers de la médaille" en essayant de comprendre le comportement des ménages par rapport à l'exploitation des ressources naturelles lorsqu'ils ont une technologie qui leur permet d'augmenter leur production ou de diminuer leurs coûts. En effet, on pourrait supposer qu'ils vont vouloir augmenter leurs surfaces cultivées, en empiétant sur la forêt, ou produire des cultures qui sont plus demandant en eau pour profiter de l'irrigation électrique. L'analyse des résultats a montré qu'il n'y a pas d'effet statistiquement significatif de l'électrification sur l'exploitation de la nappe phréatique en Inde. Ce résultat contredit la littérature sur le sujet qui prédit une surexploitation de l'eau souterraine en Inde, suite à l'utlisation de l'électricité à des fins d'irrigation [Badiani and Jessoe, 2013, Edmonds et al., 2010a]. En ce qui concerne l'effet sur la couverture forrestière, le résultat attendu était un effet positif ou neutre suite à la réduction des surfaces cultivées. Le résultat trouvé montre plutôt un effet négatif, ce qui vroudrait dire que l'électrification favorise la réduction de la couverture forrestière. Ce résultat mérite des investigations supplémentaires vue que la production électrique par les centrales à charbon génére de la pollution qui est néfaste pour la couverture forrestière. À cette limite, s'ajoute la durée couverte par cette étude qui est, à mon avis, courte (2000 à 2004). Il aurait été intéressant de réaliser cette analyse sur une plus longue période avec des données sur plusieurs années.

Le second chapitre est celui où cette thèse apporte une plus grande contribution à la littérature. Étant donné que la plupart des articles étudient théoriquement l'effet d'une politique commerciale sur la relation entre la productivité agricole et la transformation structurelle des pays en développement, ce chapitre aporte une contribution en étudiant le sujet avec une approche empirique. En suivant une méthodologie développée par Edmonds et al. [2010a] qui consiste à déterminer la perception des consommateurs, vivant au sein d'un district, sur le niveau des tariffs appliqués aux biens et services étrangers, ce chapitre considère les districts comme de petites économies ouvertes et étudie leur transformation structurelle. Comme dans le premier chapitre, l'utilisation de l'électricité à des fins d'irrigation est considérée comme la source de productivité agricole. Le principal résultat montre que suite à la baisse des tariffs, et donc à l'ouverture de l'économie aux commerce international, une expansion de l'électrification rurale entraine une réallocation de la main d'oeuvre du secteur agricole vers les secteurs non-agricoles (manufacturier, vente de détail, transport, l'éducation et le secteur publique). Ce résultat est contraire à la théorie de l'avantage comparatif de David Ricardo qui stipule qu'un pays devrait se spécialiser dans le secteur où il détient un avantage comparatif lorsqu'il s'ouvre au commerce international. Ce chapitre présente néanmoins des limites qu'il est important de souligner. La première est liée au fait d'utiliser l'année

2000 comme l'année post réforme commerciale.[75] Neuf années se sont écoulées entre les deux périodes et l'idéale aurait été d'utiliser une année plus proche de la réforme, comme 1994. Malheureusement, les données de 1994 du National Sample Survey n'inclut pas le district de résidence des individus ; ce qui rend impossible l'utilisation cette base pour étudier cette question. La seconde limite est liée à la nature endogéne de la réforme commerciale. Bien vrai que des études dans la litérature ont essayé de prouver le caractère exogéne de la réforme commerciale indienne de 1991, l'idéale aurait été de trouver un instrument pour tenir compte de l'effet des groupes de lobby sur les décisions politiques et économiques prises par les gouvernants. Une extension possible à ce chapitre serait de développer un modéle théorique qui va être calibré avec les paramétres de l'économie indienne, pour pouvoir comparer les résultats théoriques et empiriques.

Le dernier chapitre de cette thèse étudie les effets d'un choc négatif de désastre naturel sur le comportement des individus en Inde. Il regarde principalement l'effet de long terme des cyclones qui ont eu lieu durant la période de scolarité obligatoire des individus sur leur éducation et le type d'activité qu'ils occupent une fois qu'ils sont devenus de jeunes adultes. La contribution majeure qu'il apporte à la littérature est le fait qu'il étudie les impacts à long terme des catastrophes naturelles sur le capital humain, étant donné que la plupart des articles s'intéressent aux effets de long terme sur le capital physique. Il contribue également, mais dans une moindre mesure, à la littérature grandissante sur les effets de long de terme de la pandémie de COVID-19 sur l'éducation et le type d'occupation des éléves qui sont présentement affectés. Les principaux résultats trouvés par ce chapitre montrent que les cyclones peuvent engendrer un retard scolaire d'environs 13 semaines (trois mois), augmentent la probabilité d'accumuler au moins une année de retard scolaire d'environ 18 points de pourcentage, une augmentation de la probabilité de ne pas faire une éducation scolaire formelle de 4.8 points de pourcentage et une baisse de 8.1 points de pourcentage de la probabilité de finir le post-secondaire. À long terme, l'exposition durant l'enfance aux tempêtes a un impact sur le type d'activité exercée sur le marché du travail, entraînant une réduction de la probabilité d'accéder à des emplois salariés réguliers tout en augmentant la

75. la réforme commerciale a eu lieu en Inde en 1991

probabilité d'effectuer des tâches domestiques comme activité principale. La principale limite de ce chapitre est lié au fait qu'il ne tient pas compte de la migration des individus. Il fait l'hypothèse que les individus résident, en 2018, dans le district où ils ont effectué leur scolarité obligatoire. Cette hypothèse est, néanmoins, soutenue dans la littérature sur l'Inde, puisque le taux de migration inter-district est trés faible.

BIBLIOGRAPHIE

D. Acemoglu, V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi. The Network Origins of Aggregate Fluctuations. *Econometrica*, *80*(5) :1977–2016, 2012.

T. Adamopoulos. Transportation Costs, Agricultural Productivity, And Cross-Country Income Differences. *International Economic Review*, *52*(2) :489–521, 2011.

A. Adukia. Sanitation and Education. *American Economic Journal : Applied Economics*, *9*((2)) : 23–59, 2017.

H. Allcott, A. Collard-Wexler, and S. D. O'Connell. How Do Electricity Shortages Affect Industry ? Evidence From India. *American Economic Review*, *106*(3) :587–624, 2016a.

H. Allcott, A. Collard-Wexler, and S. D. O'Connell. Replication Data for : How Do Electricity Shortages Affect Industry ? Evidence from India, 2016b. URL `https://www.openicpsr.org/openicpsr/project/112983/version/V1/view`.

D. Almond, J. Currie, and V. Duque. Childhood Circumstances and Adult Outcomes : Act II. *Journal of Economic Literature*, *56 (4)*, 2018.

R. Almond and S. Sinharay. What Can Repeated Cross-Sectional Studies Tell Us About Student Growth ? . *ETS Research Report Series*, RR-12-17, 2012.

L. Anderson, C. Shinn, M. Fullilove, S. Scrimshaw, J. Fielding, J. Normand, V. Carande-Kulis, T. F. on Community Preventive Services, et al. The Effectiveness of Early Childhood Development Programs : A Systematic Review. *American Journal of Preventive Medicine*, *24*(3) :32–46, 2003.

A. Angelsen and D. Kaimowitz. Rethinking the Causes of Deforestation : Lessons From Economic Models. *The World Bank Research Observer*, *14*(1) :73–98, 1999.

A. Angelsen and D. Kaimowitz. Agricultural Technology and Tropical Deforestation. *Wallingford, UK and NY : CABI in association with the Center for International Forest Research*, page 422, 2001.

B. Archibong and F. Annan. Disease and Gender Gaps in Human Capital Investment : Evidence from Niger's 1986 Meningitis Epidemic. *American Economic Review*, *107*((5)) :530–35, 2017.

S. Asher, A. Campion, D. Gollin, and P. Novosad. The Long-run Development Impacts of Agricultural Productivity Gains : Evidence from Irrigation Canals in India. 2021.

N. Ashraf. Spousal Control and Intra-Household Decision Making : An Experimental Study in the Philippines. *American Economic Review*, *99*(4) :1245–77, 2009.

N. Ashraf, E. Field, and J. Lee. Household Bargaining and Excess Fertility : an Experimental Study in Zambia. *American Economic Review*, *104*(7) :2210–37, 2014.

O. Attanasio and V. Lechene. Tests of Income Pooling in Household Decisions. *Review of Economic Dynamics*, *5*(4) :720–748, 2002.

D. H. Autor and M. G. Duggan. The Rise in the Disability Rolls and the Decline in Unemployment. *The Quarterly Journal of Economics*, *118*(1) :157–206, 2003.

M. Azam. India's Increasing Skill Premium : Role of Demand and Supply. *The B.E. Journal of Economic Analysis and Policy*, *10*(1), 2010.

R. Badiani and K. Jessoe. The Impact of Electricity Subsidies on Groundwater Extraction and Agricultural Production. *Department of Agriculture and Resource Economics Working Paper, University of California Davis*, 2013.

L. Bakkensen and L. Barragey. Do Disasters Affect Growth ? A Macro Model-Based Perspective on the Empirical Debate. Working Paper 2016-9, 2012.

P. Balachandra. Dynamics of Rural Energy Access in India : An Assessment. *Energy*, *36*(9) : 5556–5567, 2011.

A. Banerjee, J. Suraj, and M. Kremer. Promoting School Participation in Rural Rajasthan : Results from Some Prospective Trials. 2004.

A. V. Banerjee. Educational Policy and the Economics of the Family. *Journal of Development Economics*, *74*(1) :3–32, 2004.

A. V. Banerjee, S. Cole, E. Duflo, and L. Linden. Remedying Education : Evidence from Two Randomized Experiments in India. *Quarterly Journal of Economics*, *122*(3) :1235–1264, 2007.

A. Banerji, J. Meenakshi, and G. Khanna. Social Contracts, Markets and Efficiency : Groundwater Irrigation in North India. *Journal of Development Economics*, *98*(2) :228–237, 2012.

T. J. Bartik. Who Benefits From State and Local Economic Development Policies ? 1991.

L. Beaman, E. Duflo, R. Pande, and P. Topalova. Female Leadership Raises Aspirations and Educational Attainment for Girls : A Policy Experiment in India. *Science*, *335*(6068) :582–586, 2012.

A. J. Bebbington, D. H. Bebbington, L. A. Sauls, J. Rogan, S. Agrawal, C. Gamboa, A. Imhof, K. Johnson, H. Rosa, A. Royo, et al. Resource Extraction and Infrastructure Threaten Forest Cover and Community Rights. *Proceedings of the National Academy of Sciences*, *115*(52) : 13164–13173, 2018.

G. Becker, K. Murphy, and R. Tamura. Human Capital, Fertility, and Economic Growth. *Journal of Political Economy*, *98*(5, Part 2) :S12–S37, 1990.

G. S. Becker. Investment in Human capital : A Theoretical Analysis. *Journal of Political Economy*, *70*(5, Part 2) :9–49, 1962.

M. Berlemann, M. Steinhardt, and J. Tut. Do Natural Disasters Stimulate Individual Saving ? Evidence from a Natural Experiment in a Highly Developed Country. DP No. 9026, 2015.

T. Besley and R. Burgess. Can Labor Regulation Hinder Economic Performance ? Evidence From India. *The Quarterly Journal of Economics*, *119*(1) :91–134, 2004a.

T. Besley and R. Burgess. Can Labor Regulation Hinder Economic Performance ? Evidence From India. *The Quarterly Journal of Economics*, *119*(1) :91–134, 2004b.

S. Billings, E. Gallagher, and L. Ricketts. Human Capital Investment After the Storm. mimeo, 2020.

A. Björklund and K. Salvanes. Education and Family Background : Mechanisms and Policies. In *Handbook of the Economics of Education*, volume *3*, pages 201–247. Elsevier, 2011.

M. Bjorkman Nyqvist. Income Shocks and Gender Gaps in Education : Evidence from Uganda. *Journal of Development Economics*, *105* :237–253, 2013.

P. Blaikie, T. Cannon, I. Davis, and B. Wisner. *At Risk : Natural Hazards, People's Vulnerability and Disasters*. Routledge, 2014.

O. J. Blanchard, L. F. Katz, R. E. Hall, and B. Eichengreen. Regional Evolutions. *Brookings Papers on Economic Activity*, *1992*(1) :1–75, 1992.

E. Boose, D. Foster, and M. Fluet. Hurricane Impacts to Tropical and Temperate Landscapes. *Ecological Monographs*, *64* :369–400, 1994.

E. Boose, K. Chamberlin, and D. Foster. Landscape and Regional Impacts of Hurricanes in New England. *Ecological Monographs*, *71* :27–48, 2001.

E. Boose, M. Serrano, and D. Foster. Landscape and Regional Impacts of Hurricanes in Puerto Rico. *Ecological Monographs*, *74*(2) :335–352, 2004.

J. Bound and H. J. Holzer. Demand Shifts, Population Adjustments, and Labor Market Outcomes During the 1980s. *Journal of Labor Economics*, *18*(1) :20–54, 2000.

J. Briscoe. India's Water Economy : Bracing For a Turbulent Future, Rep. 34750-IN, 102 pp., Agric. and Rural Dev. Unit. *South Asia Reg., World Bank, Washington, DC*, 2005.

J. Briscoe and R. Malik. *India's Water Economy : Bracing for a Turbulent Future*. New Delhi : Oxford University Press, 2006.

M. Browning and P.-A. Chiappori. Efficient Intra-Household Allocations : A General Characterization and Empirical Tests. *Econometrica*, pages 1241–1278, 1998.

P. Bustos, B. Caprettini, and J. Ponticelli. Agricultural Productivity and Structural Transformation : Evidence From Brazil. *American Economic Review*, *106*(6) :1320–65, 2016.

D. Card. Estimating the Return to Schooling : Progress on Some Persistent Econometric Problems. *Econometrica*, *69*(5) :1127–1160, 2001.

P. Carneiro, J. J. Heckman, and E. J. Vytlacil. Estimating Marginal Returns to Education. *American Economic Review*, *101*(6) :2754–2781, 2011.

G. Caruso. The Legacy of Natural Disasters : The Intergenerational Impact of 100 Years of Disasters in Latin America. *Journal of Development Economics*, *127* :209–233, 2017. ISSN 0304-3878. doi : https://doi.org/10.1016/j.jdeveco.2017.03.007. URL `https://www.sciencedirect.com/science/article/pii/S0304387817300317`.

E. Cavallo, S. Galiani, I. Noy, and J. Pantano. Catastrophic Natural Disasters and Economic Growth. *Review of Economics and Statistics*, *95*(5) :1549–1561, 2013.

E. A. Cavallo and I. Noy. The Economics of Natural Disasters : A Survey. Technical report, IDB, 2010.

V. Cerra and S. C. Saxena. What Caused the 1991 Currency Crisis in India ? *IMF Staff Papers*, *49*(3) :395–425, 2002.

A. K. Chakravarti. Green Revolution in India. *Annals of the Association of American Geographers*, *63*(3) :319–330, 1973.

U. Chakravorty, M. Pelli, and B. M. Ural. Does the Quality of Electricity Matter ? Evidence From Rural India. *Journal of Economic Behavior & Organization*, *107* :228–247, 2014.

H. S. R. Chan, M. L. Cropper, and K. Malik. Why Are Power Plants in India Less Efficient Than Power Plants in the United States ? *American Economic Review*, *104*(5) :586–90, 2014.

R. Chetty, J. Friedman, N. Hilger, E. Saez, D. Whitmore Schanzenbach, and D. Yagan. How Does Your Kindergarten Classroom Affect Your Earnings ? Evidence From Project STAR. *Quarterly Journal of Economics*, *126*(4) :1593–1660, 2011.

R. Chetty, N. Hendren, F. Lin, J. Majerovitz, and B. Scuderi. Childhood Environment and Gender Gaps in Adulthood. *American Economic Review Papers and Proceedings*, *106*(5) :282–88, 2016.

N. Chindarkara, Y. J. Chena, and S. Satheb. Impact of Farm Electricity Supply Management on Farm Households in Gujarat, India. 2017.

J. Colmer. Temperature, Labor Reallocation, and Industrial Production : Evidence from India. *American Economic Journal : Applied Economics*, *13*(4) :101–24, 2021.

R. Croson and G. Uri. Gender Differences in Preferences. *Journal of Economic Literature*, *47*(2) : 448–74, 2009.

J. Currie and M. Rossin-Slater. Weathering the Storm : Hurricanes and Birth Outcomes. *Journal of Health Economics*, *32*(3) :487–503, 2013.

S. Cutter, B. Boruff, and W. Shirley. Social Vulnerability to Environmental Hazards. *Social Science Quarterly*, *84*(2) :242–261, 2003.

A. De Janvry, K. Emerick, M. Gonzalez-Navarro, and E. Sadoulet. Delinking Land Rights From Land Use : Certification and Migration in Mexico. *American Economic Review*, *105*(10) :3125–49, 2015.

M. Dell, B. Jones, and B. Olken. What Do We Learn from the Weather ? The New Climate-Economy Literature. *Journal of Economic Literature*, *52*(3) :740–98, September 2014. doi : 10.1257/jel.52.3.740. URL `https://www.aeaweb.org/articles?id=10.1257/jel.52.3.740`.

C. Deppermann. Notes on the Origin and Structure of Philippine Typhoons. *Bulletin of the American Meteorological Society*, *28*(9) :399–404, 1947.

O. Dessaint and A. Matray. Do Managers Overreact to Salient Risks ? Evidence from Hurricane Strikes. *Journal of Financial Economics*, *126*(1) :97–121, 2017.

E. Deuchert and C. Felfe. The Tempest : Short- and Long-Term Consequences of a Natural Disaster for Children's Development. *European Economic Review*, *80* :280–294, 2015.

G. Di Pietro, F. Biagi, P. Costa, Z. Karpinski, and j. Mazza. The Likely Impact of COVID-19 on Education : Reflections Based on the Existing Literature and Recent International Datasets. JRC Technical Report, Joint Research Centre (JRC), European Commission, 2020.

N. K. Dubash. The Electricity-Groundwater Conundrum : Case For a Political Solution to a Political Problem. *Economic and Political Weekly*, pages 45–55, 2007.

E. Duflo. Schooling and Labor Market Consequences of School Construction in Indonesia : Evidence from an Unusual Policy Experiment. *American Economic Review*, *91*(4) :795–813, 2001.

E. Duflo. The Medium Run Effects of Educational Expansion : Evidence from a Large School Construction Program in Indonesia. *Journal of Development Economics*, *74*(1) :163–197, 2004.

E. Duflo. Women Empowerment and Economic Development. *Journal of Economic Literature*, *50* (4) :1051–1079, 2012.

C. Echevarria. Changes in Sectoral Composition Associated With Economic Growth. *International Economic Review*, pages 431–452, 1997.

E. V. Edmonds, N. Pavcnik, and P. Topalova. Trade Adjustment and Human Capital Investments : Evidence From Indian Tariff Reform. *American Economic Journal : Applied Economics*, *2*(4) : 42–75, 2010a.

E. V. Edmonds, N. Pavcnik, and P. Topalova. Replication Data for : Trade Adjustment and Human Capital Investments : Evidence from Indian Tariff Reform, 2010b. URL `https://www.openicpsr.org/openicpsr/project/113769/version/V1/view`.

J. Elsner and B. Bossak. Bayesian Analysis of U.S. Hurricane Climate. *Journal of Climate*, *14* : 4341–4350, 2001.

K. Emanuel. Increasing Destructiveness of Tropical Cyclones over the Past 30 Years. *Nature*, *436* (7051) :686–688, 2005.

K. Emanuel. Global Warming Effects on U.S. Hurricane Damage. *Weather, Climate, and Society*, *3*(4) :261–268, 2011.

K. Emanuel. Response of Global Tropical Cyclone Activity to Increasing CO2 : Results from Downscaling CMIP6 Models. *Journal of Climate*, *34* :57–70, 2021. doi : 10.1175/jcli-d-20-0367.1.

K. Emerick. Agricultural Productivity and the Sectoral Reallocation of Labor in Rural India. *Journal of Development Economics*, *135* :488–503, 2018.

E. Enarson. Violence Against Women in Disasters : A Study of Domestic Violence Programs in the United States and Canada. *Violence Against Women*, *5*(7) :742–768, 1999.

P. D. Fajgelbaum and A. K. Khandelwal. Measuring the Unequal Gains From Trade. *The Quarterly Journal of Economics*, *131*(3) :1113–1180, 2016.

S. Fankhauser and T. McDermott. Understanding the Adaptation Deficit : Why Are Poor Countries More Vulnerable to Climate Events Than Rich Countries ? *Global Environmental Change*, 27 (*1*) :9–18, 2014.

A. J. Field. Sectoral Shift in Antebellum Massachusetts : A Reconsideration. *Explorations in Economic History*, *15*(2) :146–171, 1978.

E. Field, S. Jayachandran, and R. Pande. Do Traditional Institutions Constrain Female Entrepreneurship ? A Field Experiment on Business Training in India. *American Economic Review*, *100* (2) :125–29, 2010.

R. Fishman, N. Devineni, and S. Raman. Can Improved Agricultural Water Use Efficiency Save India's Groundwater ? *Environmental Research Letters*, *10*(8) :084022, 2015.

R. M. Fishman, T. Siegfried, P. Raj, V. Modi, and U. Lall. Over-Extraction From Shallow Bedrock Versus Deep Alluvial Aquifers : Reliability Versus Sustainability Considerations for India's Groundwater Irrigation. *Water Resources Research*, *47*(6), 2011.

A. Foster and M. Rosenzweig. Inequality and the Sustainability of Agricultural Productivity Growth : Groundwater and the Green Revolution in rural India. In *Prepared for the India Policy Conference at Stanford University*, volume *5*, 2008.

A. D. Foster and M. R. Rosenzweig. Agricultural Productivity Growth, Rural Economic Diversity, and Economic Reforms : India, 1970–2000. *Economic Development and Cultural Change*, *52*(3) : 509–542, 2004.

A. D. Foster and M. R. Rosenzweig. Economic Development and the Decline of Agricultural Employment. *Handbook of Development Economics*, *4* :3051–3083, 2007.

J. A. Frankel and A. K. Rose. Is Trade Good or Bad for the Environment ? Sorting Out the Causality. *Review of Economics and Statistics*, *87*(1) :85–91, 2005.

F. A. Gabardo, J. B. Pereima, and P. Einloft. The Incorporation of Structural Change Into Growth Theory : A Historical Appraisal. *EconomiA*, *18*(3) :392–410, 2017.

P. Glewwe and M. Kremer. Schools, Teachers, and Education Outcomes in Developing Countries. volume *2* of *Handbook of the Economics of Education*, pages 945–1017. Elsevier, 2006.

P. Glewwe, M. Kremer, and S. Moulin. Many Children Left Behind ? Textbooks and Test Scores in Kenya. *American Economic Journal : Applied Economics*, *1*(1) :112–135, 2009.

P. K. Goldberg and N. Pavcnik. Distributional Effects of Globalization in Developing Countries. *Journal of Economic Literature*, *45*(1) :39–82, 2007.

P. Goldsmith-Pinkham, I. Sorkin, and H. Swift. Bartik Instruments : What, When, Why, and How. *American Economic Review*, *110*(8) :2586–2624, 2020.

D. Gollin, S. Parente, and R. Rogerson. The Role of Agriculture in Development. *American Economic Review*, *92*(2) :160–164, 2002.

D. Gollin, S. Parente, and R. Rogerson. The Food Problem and the Evolution of International Income Levels. *Journal of Monetary Economics*, *54*(4) :1230–1255, 2007.

D. Gollin, D. Lagakos, and M. E. Waugh. The Agricultural Productivity Gap. *The Quarterly Journal of Economics*, *129*(2) :939–993, 2013.

S. Goyal. Political Economy of India's Economic Reforms. *Institute for Studies in Industrial Development (ISID) Working Paper*, *4*, 1996.

J. Groen, M. Kutzbach, and A. Polivka. Storms and Jobs : The Effect of Hurricanes on Individuals' Employment and Earnings over the Long Term. *Journal of Labor Economics*, *38*(3), 2020.

V. Groppo and K. Kraehnert. The Impact of Extreme Weather Events on Education. *Journal of Population Economics*, *30* :433–472, 2017.

G. Grossman and E. Helpman. *Interest Groups and Trade Policy*. Princeton University Press, 2002.

M. D. Gupta. Informal Security Mechanisms and Population Retention in Rural India. *Economic Development and Cultural Change*, *36*(1) :101–120, 1987.

R. Gupta and M. Pelli. Electrification and Cooking Fuel Choice in Rural India. *Available at SSRN 3521473*, 2021.

J. Guryan, E. Hurst, and M. Kearney. Parental Education and Parental Time with Children. *Journal of Economic perspectives*, *22*(3) :23–46, 2008.

E. Hannum and C. Buchmann. Global Educational Expansion and Socio-Economic Development : An Assessment of Findings from the Social Sciences. *World development*, *33*(3) :333–354, 2005.

F. Hayashi and E. C. Prescott. The Depressing Effect of Agricultural Institutions on the Prewar Japanese Economy. *Journal of Political Economy*, *116*(4) :573–632, 2008.

J. Heckman. Invest in Early Childhood Development : Reduce Deficits, Strengthen the Economy. *The Heckman Equation*, *7* :1–2, 2012.

V. Henderson, A. Storeygard, and D. Weil. Measuring Economic Growth from Outer Space. *American Economic Association*, *102*, 2012a.

V. Henderson, A. Storeygard, and D. Weil. Measuring Economic Growth from Outer Space. *American Economic Association*, *102*, 2012b.

M. Henry, N. Spencer, and E. Strobl. The Impact of Tropical Storms on Households : Evidence from Panel Data on Consumption. *Oxford Bulletin of Economics and Statistics*, *82*, 2019.

B. Herrendorf, R. Rogerson, and A. Valentinyi. Growth and Structural Transformation. In *Handbook of Economic Growth*, volume *2*, pages 855–941. Elsevier, 2014.

J. Hoddinott and L. Haddad. Does Female Income Share Influence Household Expenditures ? Evidence from Côte d'Ivoire. *Oxford Bulletin of Economics and Statistics*, *57*(1) :77–96, 1995.

G. Holland. An Analytical Model of the Wind and Pressure Profiles in Hurricanes. *Monthly Weather Review*, *108* :1212–1218, 1980.

R. Hornbeck and P. Keskin. Does Agriculture Generate Local Economic Spillovers ? Short-Run and Long-Run Evidence From the Ogallala Aquifer. *American Economic Journal : Economic Policy*, *7*(2) :192–213, 2015.

S. Hsiang and A. Jina. The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth : Evidence From 6,700 Cyclones. NBER Working Papers 20352, National Bureau of Economic Research, Inc, July 2014. URL `http://ideas.repec.org/p/nbr/nberwo/20352.html`.

S. Hsu and Y. Zhongde. A Note on the Radius of Maximum Wind for Hurricanes. *Journal of Coastal Research*, *14*(2) :667–668, 1998.

H. Jain. *Trade Liberalisation, Economic Growth and Environmental Externalities : an Analysis of Indian Manufacturing Industries*. Springer, 2016.

D. Jaume and A. Willèn. The Long-Run Effects of Teacher Strikes : Evidence from Argentina. *Journal of Labor Economics*, *37*(4) :1097–1139, 2019.

S. Jayachandran and R. Pande. Why Are Indian Children So Short ? The Role of Birth Order and Son Preference. *American Economic Review*, *107*(9) :2600–2629, 2017.

R. Jensen. Economic Opportunities and Gender Differences in Human Capital : Experimental Evidence for India. Working Papers 16021, 2010.

D. Kaimowitz and J. Smith. Loss of Natural Vegetation in Brazil and Bolivia. *Agricultural Technologies and Tropical Deforestation*, page 209, 2001.

N. Kar and B. Bastia. Post-Traumatic Stress Disorder, Depression and Generalised Anxiety Disorder in Adolescents After a Natural Disaster : A Study of Comorbidity. *Clinical Practice and Epidemiology in Mental Health*, *2*(1) :17, 2006.

K. Karbownik and A. Wray. Long-Run Consequences of Exposure to Natural Disasters. *Journal of Labor economics*, *37*(3) :949–1007, 2019.

J. Kim, Y. Tong, and S. B. Sun. The Effects of Peer Parental Education on Student Achievement in Urban China : The Disparities Between Migrants and Locals. *American Educational Research Journal*, 2021.

K. Kochhar, U. Kumar, R. Rajan, A. Subramanian, and I. Tokatlidis. India's Pattern of Development : What Happened, What Follows ? *Journal of Monetary Economics*, *53*(5) :981–1019, 2006.

Z. Kone, M. Liu, A. Mattoo, C. Ozden, and S. Sharma. Internal Borders and Migration in India. *Journal of Economic Geography*, *18*(4) :729 – 759, 2018.

P. Kongsamut, S. Rebelo, and D. Xie. Beyond Balanced Growth. *The Review of Economic Studies*, *68*(4) :869–882, 2001.

M. Kuhfeld, J. Soland, B. Tarasawa, A. Johnson, E. Ruzek, and J. Liu. Projecting the Potential Impact of COVID-19 School Closures on Academic Achievement. *Educational Researcher*, *49* (8) :549–565, 2020.

P. M. Lamb. The Indian Electricity Market : Country Study and Investment Context. *Program on Energy and Sustainable Development, Center for Environmental Science and Policy, Stanford Institute for International Studies. Stanford University, Stanford, CA. http ://iis-db. stanford. edu/pubs/20975/India_Country_Study. pdf*, 2005.

M. Lindell, C. Prater, and R. Perry. *Wiley pathways introduction to emergency management.* John Wiley & Sons, 2006.

M. Lindell, C. Prater, and R. Perry. *Introduction to Emergency Management.* Wiley, 2007.

S. Maccini and D. Yang. Under the Weather : Health, Schooling, and Economic Consequences of Early-Life Rainfall. *American Economic Review*, *99*(3) :1006–1026, 2009.

K. Matsuyama. Agricultural Productivity, Comparative Advantage, and Economic Growth. *Journal of Economic Theory*, *58*(2) :317–334, 1992.

M. McMillan and K. Harttgen. What is Driving the'African Growth Miracle' ? Technical report, National Bureau of Economic Research, 2014.

C. Mellander, J. Lobo, K. Stolarick, and Z. Matheson. Night-time Light Data : a Good Proxy Measure for Economic Activity ? *PLoS ONE*, *10(10)*, 2015.

B. Min and K. Gaba. Tracking Electrification in Vietnam Using Nighttime Lights. *Remote Sensing*, *6*(10) :9511–9529, 2014.

N. Mocan and C. Cannonier. Empowering Women Through Education : Evidence from Sierra Leone. 2012.

V. Modi. Improving Electricity Services in Rural India. *CGSD Working*, *30*, 2005.

L. Monari. Power Subsidies : A Reality Check on Subsidizing Power for Irrigation in India. 2002.

K. Munshi and M. Rosenzweig. Traditional Institutions Meet the Modern World : Caste, Gender, and Schooling Choice in a Globalizing Economy. *American Economic Review*, *96*(4) :1225–1252, 2006.

K. Munshi and M. Rosenzweig. Why is Mobility in India so Low ? Social Insurance, Inequality, and Growth. Technical report, National Bureau of Economic Research, 2009.

K. M. Murphy, A. Shleifer, and R. Vishny. Income Distribution, Market Size, and Industrialization. *The Quarterly Journal of Economics*, *104*(3) :537–564, 1989.

Y. Neria, A. Nandi, and S. Galea. Post-Traumatic Stress Disorder Following Disasters : A Systematic Review. *Psychological Medicine*, *38*(4) :467, 2008.

M. Niederle and L. Vesterlund. Do Women Shy Away From Competition ? Do Men Compete Too Much ? *The Quarterly Journal of Economics*, *122*(3) :1067–1101, 2007.

D. C. North. Institutions. *Journal of Economic Perspectives*, *5*(1) :97–112, 1991.

R. Nurkse. Problems of Capital Formation in Underdeveloped Countries (New York, 1953). *NurkseProblems of Capital Formation in Underdeveloped Countries1953*, 1953.

M. Pelli and J. Tschopp. Comparative Advantage, Capital Destruction, and Hurricanes. *Journal of International Economics*, *108*(C) :315–337, 2017.

M. Pelli, J. Tschopp, N. Bezmaternykh, and K. Eklou. In the Eye of the Storm : Firms and Capital Destruction in India. Working Paper 20/203, International Monetary Fund, 2020.

C. Phelps and L. Sperry. Children and the COVID-19 Pandemic. *Psychological Trauma : Theory, Research, Practive, and Policy*, *12*(S1) :S73–S75, 2020.

R. Pielke, C. Landsea, M. Mayfield, J. Laver, and R. Pasch. Hurricanes and Global Warming. *American Meteorological Society*, pages 1571–1575, 2008.

M. M. Pitt, M. R. Rosenzweig, and M. N. Hassan. Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries. *The American Economic Review*, pages 1139–1156, 1990.

B. Platt, M. Kahn, P. Rhode, and M. Yanguas. The Effect of Natural Disasters on Economic Activity in US Counties : A Century of Data. Working Paper No. 23410, 2017.

G. Psacharopoulos. Returns to Investment in Education : A Global Update. *World development*, *22*(9) :1325–1343, 1994.

B. Ray, D. Satpathy, and K. Chandra. Evolution of National Policy on Education in India. *Vistas of Education (2013)*, 2013.

M. Rosales-Rueda. The Impact of Early Life Shocks on Human Capital Formation : Evidence from El Niño Floods in Ecuador. *Journal of Health Economics*, *62* :13–44, 2018.

P. N. Rosenstein-Rodan. Problems of Industrialisation of Eastern and South-Eastern Europe. *The Economic Journal*, *53*(210/211) :202–211, 1943.

J. P. Rud. Electricity Provision and Industrial Development : Evidence From India. *Journal of Development Economics*, *97*(2) :352–367, 2012.

N. Ryan and A. Sudarshan. Rationing the Commons. Technical report, National Bureau of Economic Research, 2020.

B. Sacerdote. When the Saints Go Marching Out : Long-Term Outcomes for Student Evacuees from Hurricanes Katrina and Rita. *American Economic Journal : Applied Economics*, *4*(1) : 109–35, 2012.

G. Santangelo. Firms and Farms : The Impact of Agricultural Productivity on the Local Indian Economy. *Unpublished Manuscript*, 2016.

P. Saxena and A. Kumar. Hydropower Development in India. *IGHEM-2010, Oct*, pages 21–23, 2010.

T. Schultz. Why Governments Should Invest More to Educate Girls. *World development*, *30*(2) : 207–225, 2002.

A. Sen, E. Hobsbawm, et al. Famine Mortality : a Study of the Bengal Famine of 1943. *Peasants in History*, 1980.

M. Shah and B. Millett Steinberg. Drought of Opportunities : Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital. *Journal of Political Economiy*, *125*, 2017.

T. Shah, J. Burke, K. G. Villholth, M. Angelica, E. Custodio, F. Daibes, J. Hoogesteger, M. Giordano, J. Girman, J. v. d. Gun, et al. Groundwater : a Global Assessment of Scale and Significance. 2007.

B. S. Sidhu, M. Kandlikar, and N. Ramankutty. Power Tariffs for Groundwater Irrigation in India : a Comparative Analysis of the Environmental, Equity, and Economic Tradeoffs. *World Development*, *128* :104836, 2020.

R. Simpson and H. Riehl. *The Hurricane and Its Impact*. Louisiana State University Press, 1981.

R. Singh. Environmental Consequences of Agricultural Development : a Case Study From the Green Revolution State of Haryana, India. *Agriculture, Ecosystems & Environment*, *82*(1-3) : 97–103, 2000.

N. Spencer, S. Polachek, and E. Strobl. How Do Hurricanes Impact Scholastic Achievement ? A Caribbean Perspective. *Natural Hazards*, *84* :1437–1462, 2016.

J. Strauss and D. Thomas. Human Resources : Empirical Modeling of Household and Family Decisions. *Handbook of Development Economics*, *3* :1883–2023, 1995.

E. Strobl. The Economic Growth Impact of Hurricanes : Evidence from U.S. Coastal Counties. *The Review of Economics and Statistics*, *93*(2) :575–589, May 2011.

Y. Takasaki. Do Natural Disasters Decrease the Gender Gap in Schooling ? *World Development*, *94(C)* :75–89, 2017.

M. Teignier. The Role of Trade in Structural Transformation. *Journal of Development Economics*, *130* :45–65, 2018.

D. Thomas. Intra-Household Resource Allocation : An Inferential Approach. *Journal of Human Resources*, pages 635–664, 1990.

W. Thornton and L. Voigt. Disaster Rape : Vulnerability of Women to Sexual Assaults During Hurricane Katrina. *Journal of Public Management & Social Policy*, 2007.

L. Tikly and A. Barrett. Social Justice, Capabilities and the Quality of Education in Low Income Countries. *International Journal of Educational Development*, *31*(1) :3–14, 2011.

T. Tombe. The Missing Food Problem : Trade, Agriculture, and International Productivity Differences. *American Economic Journal : Macroeconomics*, *7*(3) :226–58, 2015.

P. Topalova. Factor Immobility and Regional Impacts of Trade Liberalization : Evidence on Poverty From India. *American Economic Journal : Applied Economics*, *2*(4) :1–41, 2010.

P. Topalova and A. Khandelwal. Trade Liberalization and Firm Productivity : The Case of India. *Review of Economics and Statistics*, *93*(3) :995–1009, 2011.

R. Topel. Labor Markets and Economic Growth. volume *3* of *Handbook of Labor Economics*, pages 2943–2984. Elsevier, 1999.

C. Udry. Gender, Agricultural Production, and the Theory of the Household. *Journal of Political Economy*, *104*(5) :1010–1046, 1996.

A. Varshney. Mass Politics or Elite Politics ? India's Economic Reforms in Comparative Perspective. *The Journal of Policy Reform*, *2*(4) :301–335, 1998.

G. Wright. Cheap Labor and Southern Textiles Before 1880. *The Journal of Economic History*, *39*(3) :655–680, 1979.

H. Wu, M. Lindell, and C. Prater. Perceptions on Hurricane Information and Protective Action Decisions. In *Proc., of XVIII ISA World Congress of Sociology. Madrid, Spain : International Sociological Association*, 2014.

D. Yang. Coping with Disaster : The Impact of Hurricanes on International Financial Flows, 1970-2002. *The B.E. Journal of Economic Analysis and Policy*, *8*(2), 2008.

J. Yue, X. Zang, Y. Le, and Y. An. Anxiety, Depression and PTSD Among Children and Their Parent During 2019 Novel Coronavirus Disease (COVID-19) Outbreak in China. *Current Psychology*, 2020.

X. Zhou. Managing Psychological Distress in Children and Adolescents Following the COVID-19 Epidemic : A Cooperative Approach. *Psychological Trauma : Theory, Research, Practive, and Policy*, *12*(S1) :S76–S78, 2020.